

Hepatitis C Virus cell culture system

Publication number: EP1043399 (A2)

Publication date: 2000-10-11

Inventor(s): BARTENSCHLAGER RALF DR [DE] +

Applicant(s): BARTENSCHLAGER RALF DR [DE] +

Classification:






- **international:** **C07K14/02; C12N15/09; C12N15/86; C12N5/08; C12N5/10; C12N7/00; C12N7/01; A61K39/00; A61K48/00; C12R1/91; C07K14/005; C12N15/09; C12N15/86; C12N5/08; C12N5/10; C12N7/00; C12N7/01; A61K39/00; A61K48/00; (IPC1-7): A61K48/00; A61K49/00; C07K14/18; C12N15/86; C12N5/10; C12N7/01; C12N7/04**

- **European:** C07K14/02; C12N15/86; C12N7/00

Application number: EP20000105929 20000323

Priority number(s): DE19991015178 19990403

Also published as:

 EP1043399 (A3)
 EP1043399 (B1)
 EP1043399 (B9)
 DE19915178 (A1)
 US6630343 (B1)

[more >>](#)

Cited documents:

 WO9839031 (A1)
 WO9904008 (A2)
 US5851758 (A)
 US5874565 (A)
 WO9967394 (A1)

Abstract of EP 1043399 (A2)

Hepatitis C virus (HCV) cell culture system comprising human hepatoma cells that contain an integrated HCV-RNA construct (I) new. (I) contains the HCV-specific RNA segments 5'-NTR (non-translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-NTR, and a selectable (marker) gene (II). An Independent claim is also included for (I) containing the novel HCV-specific RNA segments.

.....
 Data supplied from the **espacenet** database — Worldwide



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 043 399 A2**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
11.10.2000 Patentblatt 2000/41

(21) Anmeldenummer: 00105929.4

(22) Anmeldetag: 23.03.2000

(51) Int. Cl.⁷: **C12N 15/86**, C12N 7/01,
C12N 7/04, C12N 5/10,
C07K 14/18, A61K 49/00,
A61K 48/00

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 03.04.1999 DE 19915178

(71) Anmelder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**

(72) Erfinder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**

(74) Vertreter: **Rudolph, Ulrike, Dr.
Patentanwältin
In der Schanz 10
69198 Schriesheim (DE)**

(54) **Hepatitis C Virus Zellkultursystem**

(57) Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepatomazellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

EP 1 043 399 A2

Beschreibung

[0001] Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiguriert sind.

[0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

[0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989: Science, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartenschlager 1997, Intervirology 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozessierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

[0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3' NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

[0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschrittmäßig bei allen Blutkonserven angewendet werden (müssen).

[0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Kompletsequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

[0007] Die klonierten und charakterisierten Teil- und Kompletsequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, Cell, 87,343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463; Cho et al., 1998, J. Biol. Chem., 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

[0008] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

[0009] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können, (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, Proceedings of the National Academy of Sciences, USA, 90, 6037-6041; Mizutani et al. 1996, Journal of Virology, 70, 7219-7223; M. Ikeda et al. 1998, Virus Res. 56, 157; Fournier et al. 1998, J. Gen. Virol. 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, Journal of General Virology, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Proteine in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

[0010] Desweiteren ist aus den Publikationen von Yoo et al. (1995, Journal of Virology, 69, 32-38) und von Dash et al., (1997, American Journal of Pathology, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels in vitro Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeführten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.

[0011] Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines HCV- Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.

[0012] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das im

wesentlichen die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

[0013] Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

[0014] Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

[0015] Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markergen (Selektionsgen) umfaßt.

[0016] Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

[0017] Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenese und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

[0018] Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

[0019] Eine Ausführungsform des HCV-RNA-Konstrukts, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO:1 umfaßt.

Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO:2 oder SEQ ID NO:3 oder SEQ ID NO:4 oder SEQ ID NO:5 oder SEQ ID NO:6 oder SEQ ID NO:7 oder SEQ ID NO:8 oder SEQ ID NO:9 oder SEQ ID NO:10 oder SEQ ID NO:11 umfassen.

[0020] Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3' NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

50

55

- 5 (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA
GCTTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC
AGACTGGAGT
- 10 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCAGACT GGAGC
- 15 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTAATCTTT TTTTTTTTTTCT TTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCAGACT GCAGC
- 20 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT
- 25 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT
- 30 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTT TTGAGAGAGA GAGTCTCACT
CTGTTGCCCA GACTGGAGT
- 35 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGA CTGGAGT
- 40 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTTAAT CTTTTTTTTT TTTTCCTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 45 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAATC TTTTTTTTTT TTTCTTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

50 Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blasticidin.

[0021] Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

[0022] Eine Alternative zu den Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-

Selektion durchgeführt werden kann.

[0023] Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5' NTR, d.h. strangabwärts der 5' NTR bzw. strangaufwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion im Bereich der 3' NTR oder an anderer Stelle des HCV-Genoms oder -Subgenoms, z.B. innerhalb des Polypeptids.

[0024] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

[0025] Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des inklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0026] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

[0027] Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

[0028] Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

[0029] Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasen) oder dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

[0030] Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

[0031] Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms inkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polypeptids oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'high-throughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.

- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflusst werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- 5 • Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder
10 anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausch/s in das Ursprungs-konstrukt kann deren
15 Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von
20 Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellkultursystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb in der Lage infektiöse Viren zu produzieren.

25 **[0033]** Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:

- Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen: Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertio-
30 nen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringerter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
- Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als leberzellspezifische Genfährten in der Gentherapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte
35 herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
- Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das aller-
45 dings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumselektion unterzogen oder, in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine
50 Antibiotikumresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.

55 **[0034]** Das erfindungsgemäße Zellkultursystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeführt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts führen, sind dem Fachmann als adaptive

Mutationen bekannt. Die Erfindung umfaßt deshalb auch Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Sie umfaßt desweiteren ein Verfahren zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, sowie zellkultur-adaptierte Mutanten von HCV-RNA-Konstrukten, HCV-Vollängengenomen und HCV-Teilgenomen mit im Vergleich zu den ursprünglichen Konstrukten, Teil- oder Vollängengenomen erhöhter Replikationseffizienz.

[0035] Das erfindungsgemäße Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, ist dadurch gekennzeichnet, daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 19 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

[0036] Bei einer vorteilhaften Weiterbildung dieses Herstellungsverfahrens werden die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 eingeschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone geerntet und daraus die HCV-RNA-Konstrukte isoliert.

Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

[0037] Das erfindungsgemäße Verfahren zur Herstellung von Mutanten eines HCV - Vollängengenoms oder eines HCV- Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV- Vollängengenom oder - Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz zeichnet sich dadurch aus, daß man mit Hilfe eines der beiden vorstehend genannten Herstellungsverfahren eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt, diese aus den Zellen isoliert, mit im Stand der Technik bekannten Methoden kloniert und sequenziert und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Mutationen bestimmt, und diese Mutationen dann entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, welche die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängen- oder -teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Markergen die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

[0038] Die erfindungsgemäßen zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotid- und/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

[0039] Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nicht-selektierbaren Reportergen (z.B. Luziferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektionierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

[0040] Eine Gruppe ganz bevorzugter HCV-RNA-Konstrukte, HCV-Vollängengenome und HCV-Teilgenomen mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist

EP 1 043 399 A2

dadurch gekennzeichnet, daß sie einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der folgenden Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

5

Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:

SEQ ID-NO: 1

10 **[0041]**

Name: l389/Core-3'/wt

Aufbau (Nukleotidpositionen):

15

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

20

4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7

25

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)

30

14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 2

35 **[0042]**

Name: l337/NS2-3'/wt

Aufbau (Nukleotidpositionen):

40

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

45

4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8407-8637: HCV 3' nicht-translatierte Region

55

SEQ ID-NO: 3

[0043]

- 5 Name: l389/NS3-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
 5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 - 15 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
 8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
 9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 10. 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 4

[0044]

- 25 Name: l337/NS3-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 30 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
 5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 - 35 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3859-4641: Nichtstrukturprotein 4B (NS4B)
 8. 4642-5982: Nichtstrukturprotein 5A (NS5A)
 9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 10. 7759-7989: HCV 3' nicht-translatierte Region

SEQ ID-NO: 5

[0045]

- 45 Name: l389/NS2-3'/wt
Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 50 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
 5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
 - 55 6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
 9. 5302-6642: Nichtstrukturprotein 5A (NS5A)

EP 1 043 399 A2

10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

5

[0046]

Name: I389/NS3-3'/9-13F

Aufbau (Nukleotidpositionen):

10

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 15 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

20

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

25 SEQ ID-NO: 7

[0047]

Name: I389/Core-3'/9-13F

Aufbau (Nukleotidpositionen):

30

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 35 4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F

40

45

5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

50 SEQ ID-NO: 8

[0048]

Name: I389/NS3-3'/5.1

Aufbau (Nukleotidpositionen):

55

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

EP 1 043 399 A2

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase

6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor

7. 3871-4653: Nichtstrukturprotein 4B (NS4B)

8. 4654-5994: Nichtstrukturprotein 5B (NS5A)

10

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 9

15 [0049]

Name: l389/Core-3'/5.1

Aufbau (Nukleotidpositionen):

20

1. 1-341: HCV 5' nicht-translatierte Region

2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

25

5. 1813-2385: HCV Core Protein; Strukturprotein

6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein

7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein

8. 4051-4239: Protein p7

30

9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease

10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase

11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor

12. 6946-7728: Nichtstrukturprotein 4B (NS4B)

13. 7729-9069: Nichtstrukturprotein 5A (NS5A)

35

14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase

15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 10

40 [0050]

Name: l389/NS3-3'/19

Aufbau (Nukleotidpositionen):

45

1. 1-341: HCV 5' nicht-translatierte Region

2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters

4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

50

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase

6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor

7. 3871-4653: Nichtstrukturprotein 4B (NS4B)

55

8. 4654-5994: Nichtstrukturprotein 5A (NS5A)

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 11

[0051]

5 Name: I389/Core-3'/19

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
5. 1813-2385: HCV Core Protein; Strukturprotein
- 15 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 20 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

25

[0052] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

30 **Fig. 1 A:** Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

35 Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo^R), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck (Δ) markiert.

45 **Fig. 1 B:** Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.

50 **Fig. 1 C** : Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone. Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.

55 **Fig. 2 A:** Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I₃₇₇/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I₃₇₇/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 μ g 9-13 DNA vor der PCR (zwecks Ausschluß

eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem µg Gesamt-DNA des Zellkons 9-13 erhalten wurde.

- 5 **Fig. 2 B:** Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus" -Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrang" und "Plusstrang" bezeichnen die Spezifität der radioaktiven RNA-Sonden.
- 10 **Fig. 2 C:** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.
- 15 **Fig. 3 A:** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.
- 20 **Fig. 3 B:** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.
- 25 **Fig. 4:** Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.
- 30 **Fig. 5:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Antibiotikumresistenzgen (A) innerhalb der für das Polypeptid kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3' NTR (bicistronische RNA).
- 35 **Fig. 6:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reportergen wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polypeptid gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schraffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (D) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).
- 40 **Fig. 7:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.
- 45 **Fig. 8:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.
- 50 **Fig. 9:** Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in

naïve Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1 beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphorimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den in vitro Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 µg eines Plasmids zugesetzt, das die Expression der Luziferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luziferaseaktivität im jeweiligen Zellysat bestimmt. Die Anzahl der Kolonien wird jeweils auf die Luziferaseexpression normiert.

Fig. 10: Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukts I377/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den reklonierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* I verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.

Fig. 11 A: Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt I₃₈₉/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luziferasegen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NS5B RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angedeutet. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine in vitro Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naive Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.

Fig. 11 B: Vergleich der Luziferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt I₃₈₉/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-13F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luziferaseaktivitäten luminometrisch bestimmt.

Fig. 12: Selektionierbare HCV-Vollängengenome (Konstrukte I₃₈₉/core-3'/5.1 und I₃₈₉/core-3'/9-13F).

(A) Schematische Darstellung des Vollängenkonstrukts. Der Bereich zwischen den beiden ange-deuteten Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hoch-adaptierten RNA-Varianten 5.1. oder 9-13F.

(B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 µg in vitro transkribierter RNA der unter A dargestellten Konstrukte I₃₈₉/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experimentes.

(C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden in vitro Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das *neo*-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10⁸ Molekülen der angegebenen in vitro Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 µg Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von in vitro transkribierter I₃₈₉/core-3'/5.1-RNA bzw. I₃₈₉/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

Fig. 13: HCV-RNA-Konstrukte mit einem Reporter-gen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reporter-gen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reporter-genprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reporter-genprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.

Fig. 14: Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.

Fig. 15: Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

Beispiel 1: Herstellung von HCV-RNA-Konstrukten

(A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

[0053] Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h. die HCV-RNA wie nachfolgend beschrieben isoliert:

[0054] Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGGCCTATTGG CCTGGAG) und dem 'expand reverse transcriptase'-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR=polymerase chain reaction) durchgeführt, und zwar unter Verwendung des 'expand long template'-Systems (Boehringer Mannheim, Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACAGCCCGCGTCATAGCACTCG) oder S4542 und A9386 (TTAGCTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten 'X-tails' (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor inseriert. Abweichungen von der Konsensussequenz wurden mittels ortsgerichteter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3' Ende der 3' NTR (Kolykhalov et al. ibid.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch

gefunden.

[0055] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

(B) Synthese selektierbarer HCV-Subgenom-Konstrukte

5

[0056] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, *Molecularcloning: a laboratory manual*, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, *Current Protocols in Molecular Biology*, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5' NTR inseriert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebenso gut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstrukts inseriert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Enzephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0057] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt: Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kofusionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0058] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

[0059] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0060] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

(C) Synthese selektierbarer HCV-Genom-Konstrukte

[0061] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit NcoI und SpeI restringiert und mittels präparativer Agarosegelelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem NcoI/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/SpeI-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit PmeI und SpeI restringiert und mit dem analog restringierten I₃₈₉/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

(D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Transkripten

[0062] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Transkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM MgCl₂, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/µl, 50 µg/ml restringierte DNA und ca. 2 Einheiten/µl T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, *Bio Techniques* 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro µg DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittel optischer Dichtemessungen quantifiziert und

ihre Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7

5 **[0063]** Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

10 **[0064]** Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, *Cancer Res.* **42**, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8 x 10⁶ Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

15 **[0065]** Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klone 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.

20 **[0066]** Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

25 **[0067]** Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, *Anal. Biochem.* **162**, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von 0,5 - 1 x 10⁶ Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 10⁹ in-vitro-Transkripte (ivtr.), die zu den I₃₈₉/NS2-3'/wt oder den I₃₈₉/NS3-3'/wt Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.

40 **[0068]** Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40µg/ml, 1h, 37°C) in 10mMTris, pH7,5, 1mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend 4 - 8 x 10⁴ Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTTCGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10⁷ Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.

Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrol-

len, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

[0069] Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren: Es wurden jeweils 10^6 - 10^2 Plasmid Moleküle ($I_{377}/NS3-3'/wt$) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils $1 \mu g$ 9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem μg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder $2 \mu g$ Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrang', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrang', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca. 10^3 Kopien/ μg Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [3H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [3H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5×10^5 Zellen der Klone 9-13 und 8-1 mit $100 \mu Ci$ [3H]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von $4 \mu g/ml$ Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 und Lohmann et al., 1997, *J Virol.* 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflusst worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Reklonierung der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der reklonierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transkribierten Transkript des HCV-Konstrukts $I_{377}/NS3-3'/wt$ übereinstimmt.

[0070] Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [^{35}S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nicht-denaturierenden Bedingungen (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519) und unter Verwendung von drei verschiedenen Antiseren (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abge-

trennt worden.. Die Immunokomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonstrukt I₃₇₇/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden.

5 Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum (3/4) bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0071] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transients Expression desselben bicistronischen HCV-RNA-Konstrukts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunofluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antiseren durchgeführt (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azeton fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antiseren inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegengefärbt.

[0072] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antiseren war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antiseren führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen.

20 Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet — so wie das bei den meisten RNA-Viren der Fall ist.

[0073] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellkultursystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

[0074] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektioniert, wobei hier jedoch selektionierbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [³H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunpräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukts, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.

[0075] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das einklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende

Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

5

Beispiel 5: Kotransfektion eines HCV-RNA-Konstrukts mit einem separaten Luziferase-Transfektionskonstrukt

[0076] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen ver-
 mittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit
 einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden
 ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen,
 insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen.
 Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der
 Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das
 ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt,
 die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird
 das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestim-
 mung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann
 ebenso gut ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression
 von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es
 kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als
 sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmark-
 ers ein Maß für die Replikation der viralen DNA.

25

Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Gentherapien

[0077] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementie-
 rende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispiels-
 weise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen
 sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die
 Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also
 virus-ähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des inkorporierten Fremd-
 gens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Kon-
 strukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäße HCV-
 Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vor-
 genannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-
 komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als
 Expressionsvektor.

40

Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte

(A) Isolationsverfahren

45

[0078] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte
 wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben
 transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in
 diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-
 RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt
 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B
 beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca. 10^9 Moleküle HCV-RNA enthielt, wurde anschließend
 per Elektroporation in naive Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden 10^9 in vitro Transkripte der ana-
 logen neo-HCV-RNA, die mit isolierter Gesamt-RNA aus naiven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg
 aufgefüllt worden war, in naive Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien,
 ausgedrückt in 'colony forming units (cfu) pro Mikrogramm RNA' in den beiden Ansätzen bestimmt. Bei einer Konzen-
 tration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der *isolierten* Gesamt-
 RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit dersel-

55

ben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

[0079] Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplifikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser reklonierten RNAs mit der Sequenz der RNA, die ursprünglich in die naiven Huh-7 Zellen eingeschleust wurde ergab, daß die reklonierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10). *Sfi*I-Fragmente dieser reklonierten Mutanten wurden im Austausch gegen das analoge *Sfi*I-Fragment des ursprünglichen Replikonkonstrukts in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naive Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt. Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der reklonierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten. Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.

[0080] Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naive Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.

Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den *neo*-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektionieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.

Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naiven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt I₃₈₉/NS3-3' erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IRES besaß (I₃₇₇/NS3-3'; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naive Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naive Huh-7 Zellen transfiziert und analog selektioniert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die *neo*-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfi*I verdaut und in das *Sfi*I-restringierte Ausgangskonstrukt I₃₈₉/NS3-3' inseriert. Insgesamt wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naive Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten *neo*-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. "adaptiven Mutationen") mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNA.

(B) Modifikationsverfahren

[0081] Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukts. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht

werden können, die kein selektierbares Markergen mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffizienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprechen. Zwecks einfacher Messung wurde das *neo*-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NS5B RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

15 **Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome**

[0082] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, der die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das SfiI-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transferiert (Fig. 12). Dieses HCV-Genom wurde in naive Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NS5B RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängengenom enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierten HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfindungsgemäßen System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfindungsgemäßen HCV-Zellkultursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.

[0083] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Beispiel 10: Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfährer für die Gentherapie oder als Expressionsvektor.

[0084] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsidiert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

Beispiel 11: Herstellung von monocistronischen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0085] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Tabelle 1

Spezifische Infektiositäten (cfu/ μ g RNA) der HCV-RNA-Konstrukte mit adaptiven Mutationen, die bei der 9-13F Mutante gefunden und in das parentale HCV-RNA-Konstrukt 1 ₃₈₉ /NS3-3'/wt eingeführt wurden		
Aminosäureaustausch ¹	HCV-Protein	cfu/ μ g RNA ²
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

¹ Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

² Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/ μ g RNA) des parentalen HCV-RNA-Konstrukts I ₃₈₉ /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte RNA-Variante	cfu/ μ g RNA ¹	
Wildtyp	30 - 50	
9-13 C	100 - 1.000	
9-13 I	100 - 1.000	
9-13 F	1.000 - 10.000	
5.1	50.000 - 100.000	
19	50.000 - 100.000	

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 3: Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-RNA-Konstrukt I₃₈₉/NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19

HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
	9346	T > C	Val > Ala
9-13 F	3866	C > T	-
	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
	8486	C > T	-
	8991	C > G	Arg > Gly
NK5.1	4180	C > T	Thr > Ile
	4679	C > T	-

5		4682	T > C	-
		5610	C > A	Leu > Ile
		6437	A > G	-
10		6666	A > G	Asn > Asp
		6842	C > T	-
		6926	C > T	-
15		6930	T > C	Ser > Pro
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
20	NK19	3946	A > G	Glu > Gly
		4078	C > G	Ala > Gly
		4180	C > T	Thr > Ile
25		4682	T > C	-
		5610	C > A	Leu > Ile
		5958	A > T	Met > Leu
30		6170	T > A	-
		6596	G > A	-
35		6598	C > G	Ala > Gly
		6833	C > T	-
		6842	C > T	-
40		6930	T > C	Ser > Pro
		7141	A > G	Glu > Gly
		7320	C > T	Pro > Ser
45		7389	A > G	Lys > Glu
		7735	G > A	Ser > Asn

50

[0036] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen der zellkulturadaptierten HCV-RNAs. Die Zahlen
55 beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

SEQUENCE LISTING

5 <110> Bartenschlager, Ralf
 <120> Hepatitis C Virus cell culture system
 <130> ba-1
 <140> 199 15 178.4
 10 <141> 1999-04-03
 <160> 11
 <170> PatentIn Ver. 2.1
 <210> 1
 15 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 <400> 1
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 20 tcttcacgca gaaagcgctc agccatggcg ttagtatgag tgctgtgcag cctccaggac 120
 ccccctccc gggagagcca tagtggctcg cggaaccggg gactacaccg gaattgccag 180
 gacgaccggg tctcttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaaagaaa aaccaaaccgt aacaccaacg ggcgcgcat gattgaacaa gatggattgc 420
 acgcaggttc tccgcccgtc tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 25 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgactgttgc actgaagcgg 660
 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctctcgccga gaaagtatcc atcatggctg atgcaatgcg gcgctgcat acgcttgatc 780
 cggtactctg ccattcgcac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 30 tggaaagccg tcttgcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgcag 900
 ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgacc 960
 atggcgatgc ctgcttgcg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggccg gctgggtgtg gcgacccgct atcaggacat agcgttggct acccgtaga 1080
 ttgtgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatc gcagcgcatc gccttctatc gccttcttga cgagtcttct tgagtttaaa 1200
 cagaccacaa cgttttccct ctacgggatc caattccgcc cctctccctc cccccccct 1260
 35 aacgttactg gccgaagccg cttggaataa ggcgggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcgctcttt tggcaatgtg agggcccggg aacctggccc tgtcttcttg 1380
 acgagcattc ctagggtgtc ttccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
 gtgaaggag cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcagacctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
 40 gaaagagtca aatggtctc ctcaagcgtc ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccatt gtatgggac tgatctggg cctcgggtgca catgctttac atgtgttag 1740
 tgcaggttaa aaaacgtcta ggcctccga accacgggga cgtggttttc ctttgaaaaa 1800
 cacgataata ccattggcac gaatcctaaa cctcaagaa aaaccaaacy taacaccaac 1860
 cgccgccac aggacgtcaa gttcccggc ggtgggtcaga tcgtcgggtg agtttacctg 1920
 ttgcgcgca ggggcccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
 caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
 45 gctcagcccg ggtaccctg gccctctat ggcaatgagg gcttgggggtg ggcaggatgg 2100
 ctctgtcac ccgtggctc tcggcctagt tggggcccca cggacccccc gcgtagggtcg 2160
 cgcaatttgg gtaagggtcat cgataccctc acgtgcggct tcgcgatct catggggtac 2220
 attcgcgtcg tcggcgcccc cctagggggc gctgccaggg ccctggcgca tggcgctccg 2280
 gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc ctttctatc 2340
 ttcttttggt ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
 50 gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
 gcggacatga tcatgcatac ccccggtgc gtgcctcgc ttccggagaa caactcctcc 2520
 cgctgctggg tagcgtcac tcccacgtc gcggccagga acgctagcgt cccactacg 2580
 acgatacgac gccatgtcga tttgctcgtt gggcgcgctg ctctctgctc cgctatgtac 2640
 gtgggagatc tctcggtatc tgttttctc gtgcgccagc tgttcacctt ctgcctcgc 2700

55

EP 1 043 399 A2

5 cggcaccgaga cagtagacagga ctgcaattgc tcaatataatc ccggccacgt gacaggtcac 2760
 cgtatggcctt gggatattgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
 ttaactccgga tcccacaagc tgtcgtggat atggtggcgg gggccatttg gggagtccca 2880
 gcggggccttg cctactatcc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
 ctctttgccc gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
 ctcgggatta cgtccctctt ttcacccggg tcatccaga aaatccagct tgtaaacacc 3060
 aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120
 ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
 gccagctgca gcccacatga cgtcttcgct cagggtggg gggccatcac ttacaatgag 3240
 10 tcacacagct cggaccagag gccctattgt tggcactacg caccocggcc gtgcggtatc 3300
 gtaccccgcg cgcagggtgtg tgggtccagt tactgcttca ccccaagccc tgtcgtgggtg 3360
 gggacgaccg accgggtcgg cgtccctacg tacagtggg gggagaatga gacggacgtg 3420
 ctgcttctta acaacacgcg gccgcgcgaa ggcaactggt ttgctgttac atggatgaat 3480
 agcactgggt tcaccaagac gtgcgggggc ccccggtgta acatcgggg gatcggcaat 3540
 aaaacctga cctgccccac ggactgcttc cggaaagcac ccgagggcac ttacaccaag 3600
 tgtggttcga ggccttggtt gacacccaga tgcctggctc actacccata caggcttttg 3660
 15 cactaccctt gcaactgtcaa ctttaccatc ttcaaggta ggaatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgttgtaa cctggaggac 3780
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtggca ggtattgcc 3840
 tgttccctca ccaccctacc ggctctgtcc actggttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
 gagtatgtcc tgttgcctct ccttctctg gcggaocgcg ccgtctgtgc ctgcttgtgg 4020
 atgatgctgc tgatagctca agctgaggcc gccctagaga acctgggtgt cctcaacgcg 4080
 20 gcacccgttg ccggggcgca tggcaattctc tccctcctcg tgttctctg tgcgtcctgg 4140
 tacatcaagg gcaggctggt cccctggggcg gcataatgccc tctacggcgt atggccgcta 4200
 ctctgtctcc tgcctggcgtt accaccacga gcatacggca tggaccggga gatggcagca 4260
 tctgtcggag gcgcgggttt cgttaggtctg atactcttga ccttgtcacc gcaactataag 4320
 ctgttctctg ctaggctcat atggttggtta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggcacccccc cctcaacgtt cggggggggc gcgatgcctg catcctcctc 4440
 25 acgtgcgcga tccacccaga gctaactttt accatcacca aaatcttgtc cgccatactc 4500
 ggtccactca tgggtgctcca ggctggtata accaaagtgc cgtactctgt gcgcgcacac 4560
 gggctcattc gtgcattgat gctggtgcgg aaggttgcgt ggggtcatta tgtccaaatg 4620
 gctctcatga agttggcggc actgacaggt acgtacgttt atgacctct caccocactg 4680
 cgggactggg cccacgcggg cctacgagac cttgcggtgg cagttgagcc cgtcgtcttc 4740
 tctgatattg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
 atcttgggccc tgcccgctcc cgcgcgaggg gggagggaga tacatctggg accggcagac 4860
 30 agccttgaag ggcagggttg gcgactcctc gcgcctatta cggcctactc ccaacagacg 4920
 cgaggcctac ttggtgcat catcactagc ctcacaggcc gggacaggaa ccaggctcag 4980
 ggggaggtcc aagtgtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
 gtgtgttgga ctgtctatca tgggtccggc tcaaagaccc ttgcgggccc aaaggggcca 5100
 atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
 gcgcgttccct tgacaccatg cacctgcggc agctcggacc gtagcctact ctccccagg 5220
 35 gccatgttca ttcgggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5280
 cccgtctcct acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcacgct 5340
 gtgggcatct ttcgggtgcg cgtgtgcacc cgagggttg cgaaggcggt ggactttgta 5400
 cccgtcgagt ctatggaac cactatgcgg tcccgggtct tcacggacaa ctgctcccct 5460
 ccggccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
 agcactaagg tgcgggtgcg gtatgcagcc caagggtata aggtgcttgt cctgaacctg 5580
 tccgtcggcc ccaccctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640
 40 aacatcagaa ccggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
 ggcaagtttc ttgcgcaggg tgggtgtctc gggggcgcc atgacatcat aatatgtgat 5760
 gagtgccact caactgactc gaccatctc ctgggcatcg gcacagtcct ggaccaagcg 5820
 gagacggctg gagcgcgact cgtcgtgctc gccaccgcta ccgctccggg atcggtcacc 5880
 gtgccacatc caaacatcga ggaggtggtc ctgtccagca ctggagaaat ccccttttat 5940
 ggcaaaagcca tcccacatga gaccatcaag ggggggagge acctcatttt ctgccattcc 6000
 45 aagaagaaat gtgatgagct cgcgcggaag ctgtccggcc tcggactcaa tgtcgtagca 6060
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgtcgtagca 6120
 acggacgctc taatgacggg ctttaccggc gatttcgact cagtgtatga ctgcaatata 6180
 tgtgtcaccg agacagtcga cttcagcctg gaccgcacct tcaccattga gacgacgacc 6240
 gtgccacaag acgcggtgtc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcattgtcga ttcctcggtt 6360
 ctgtgcgagt gctatgacgc gggctgtgct tggtaacgagc tcacgcccgc cgagacctca 6420
 50 gttaggttgc gggcttacct aaaaacacca ggggtgcccg tctgcccagg ccatctggag 6480
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540
 aagcaggcag gagacaactt cccctacctg gtagcatacc aggcctacggt gtgcgccagg 6600
 gctcaggctc caccctccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660

55

EP 1 043 399 A2

	acgctgcacg	ggccaacgcc	cctgctgtat	aggctgggag	ccgttcaaaa	cgaggttact	6720
	accacacacc	ccataaccaa	atacatcatg	gcatgcatgt	cggctgacct	ggaggctcgtc	6780
	acgagcacct	gggtgctggg	aggcggagtc	ctagcagctc	tggcccgcta	ttgctcgaca	6840
5	acaggcagcg	tggtcattgt	gggcaggatc	atcttgtccg	gaaagccggc	catcattccc	6900
	gacagggaag	tcctttaccg	ggagttcgat	gagatggaag	agtgcgcctc	acacctccct	6960
	tacatcgaac	agggaatgca	gctcgccgaa	caattcaaac	agaaggcaat	cgggttgctg	7020
	caaacagcca	ccaagcaagc	ggaggctgct	gctcccgtgg	tggaatccaa	gtggcggacc	7080
	ctcgaagcct	tctgggcgaa	gcatatgtgg	aatttcatca	gcgggataca	atattttagca	7140
	ggcttgtcca	ctctgcctgg	caaccocgg	atagcatcac	tgatggcatt	cacagcctct	7200
	atcaccagcc	cgtcaccac	ccaacatacc	ctcctgttta	acatcctggg	gggatgggtg	7260
10	gccgcccaac	ttgctcctcc	cagcgtgctc	tctgctttcg	taggcgcggg	catcgctgga	7320
	gcggctgttg	gcagcatagg	ccttgggaag	gtgcttggg	atattttggc	aggttatgga	7380
	gcagggttg	cagggcgcgt	cgtggccttt	aaggctcatg	gcggcgagat	gccctccacc	7440
	gaggacctgg	taaacctact	ccctgctatc	ctctcccctg	gcgccctagt	cgtcggggtc	7500
	gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgctt	7620
15	gagagcgacg	ctgcagcacg	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
	ctgaagaggg	ttcaccagtg	gatcaacgag	gactgtctca	cgccatgctc	cggctcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gcggggagtc	ccctttctct	catgtcaacg	tggttacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taaacagttg	7980
20	catggaacat	ttcccattaa	cgcgtacacc	acggggccct	gcacgccctc	cccggcgcca	8040
	aattattcta	gggcgctgtg	gcgggtggct	gctgaggagt	acgtggaggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtgcct	gtgtcagggt	8160
	ccggcccccgc	aattcttcac	agaagtggat	gggggtcggt	tgacacagga	cgctccagcg	8220
	tgcaaacccc	tcctacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctgggt	8280
	gggtcacacg	ttccatgcga	gcccgaaccg	gacgtagcag	tgctcacttc	catgctcacc	8340
	gacctctccc	acattacggc	ggagacggct	aagcgtaggc	tgcccagggg	atctccccc	8400
25	tccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actcccggga	cgtcgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atgggcggga	acatcacccg	cgtggagtcg	gaaaataagg	tagtaatttt	ggactcttct	8580
	gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
	aggtccagga	aattccctcg	agcgtatccc	atatgggcac	gcccggatta	caacctcca	8700
	ctggttagagt	cctggaagga	cccggactac	gtccctccag	tggtacacgg	gtgtccattg	8760
	cgccctgcca	aggccctccc	gataccacct	ccacggagga	agaggacggt	tgtcctgtca	8820
30	gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tcgtcgcccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	ttcatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggctc	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgctgct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgttgcggag	9120
	gaaaccaagc	tgcccatcaa	tgcaactgagc	aactctttgc	tcctgtacca	caacttggtc	9180
35	tatgtacaaa	catctcgcag	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
	caggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaagc	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgcgt	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggcaa	aaaatgaggt	ttctctgcgc	caaccagaga	agggggggcg	caagccagct	9540
	cgcttatctg	tattcccaga	tttgggggtt	cgtgtgtgct	agaaaaatgg	cctttacgat	9600
40	gtggtctcca	ccctccctca	ggcgtgatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tggaagcgca	agaaatgccc	tatgggcttc	9720
	gcatatgaca	cccgctgttt	tgactcaacg	gtcactgaga	atgacatccg	tggtgaggag	9780
	tcaatctacc	aatgttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtccaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaaact	cggctatcgc	9900
	cgggtccgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
45	aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
	gacgaccttg	tcgttatctg	tgaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggaccgcgc	caaaccagaa	10140
	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
	ggcaaaaagg	tgtactatct	cacccgtgac	cccaccaccc	cccttgccgcg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggttaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catctctctt	ccatccttct	agctcaggaa	10380
50	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctacctc	agatcattca	acgaactccat	ggccttagcg	cattttcaact	ccatagttac	10500
	tctccagggt	agatcaatag	ggtggcttca	tgccctcagga	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgtaggcc	tactgtccca	gggggggagg	10620

55

EP 1 043 399 A2

	gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccg	ctgcgtecca	gttggattta	tccagctggt	tcgttgctgg	ttacagcg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcacgtg	gtgcctactc	10800
5	ctactttctg	taggggtagg	catctatcta	ctccccaacc	gatgaacggg	gagctaaaca	10860
	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
	tttttttttt	tttttttttc	tccttttttt	ttcctctttt	tttccttttc	tttccttttg	10980
	tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

10

15

20

25

30

35

40

45

50

55

<210> 2
 <211> 8637
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 2

10

15

20

25

30

35

40

45

50

55

```

gccagccccc gattgggggc gacactccac catagatcac tccccgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctgtgag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gactacacgg gaattgccag 180
gacgaccggg tctttctctg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttggtgc gcgaaaggcc ttgtgttact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcagc aatcctaaac 360
ctcaaaagaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttcocg ctgtcagcgc agggcgcccc ggttcttttt gtcaagaccg 540
acctgtccgg tcacctgaat gaactgcagg acgaggcagc gcgctatcg tggctggcca 600
cgacggggct tccctgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcog gggcaggatc tcctgtcac tcaccttgc cctgccgaga 720
aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaaa catcgcatcg agcgagcac tactcggatg gaagccggtc 840
ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900
ccaggctcaa ggcgcgcatg ccgcagcgcg aggatctcgt cgtgacccat ggcgatgcct 960
gcttgccgaa tatcatgggt gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020
tggtgtgggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080
ttggcgcgca atgggtcgac cgttctctcg tcttttacgg tatcgccgct cccgattcgc 1140
agcgcatcgc cttctatcgc cttcttgacg agttctctg agtttaaaac gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tggaaataag ccggtgtgcg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcgggaaa cctggccctg tcttcttgac gagcattcct 1380
atggggtctt ccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
gttctcttgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500
aacccccacg ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaaggcg cacaaccoca gtgccacgtt gtgagttgga tagttgtgga aagagtcaa 1620
tggtctctct caagcgtatt caacaagggg ctgaaggatg cccagaaggc accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgttttagt gaggttaaaa 1740
aacgtctagg cccccgaac cagggggacg tggttttcct ttgaaaaaca cgataatacc 1800
atggaccggc agatggcagc atcgtgcgga ggccgggttt tcgtaggctc gatactcttg 1860
acctgttcac cgcactataa gctgttctc gctaggctca tatggtggtt acaatatttt 1920
atcacccagg ccgaggcaca cttgcaagtg tggatcccc cctcaacgt tcgggggggc 1980
cgcatgcccg tcactctcct cactgtcgcg atccaccag agctaatttt taccatcac 2040
aaaatcttgc tcgccatact cggctccact atggtgtctc aggtgtgtat aaccaaagt 2100
ccgtactctg tcgcgcgaca cgggtctatt cgtgcacgca tgctgtgtcg gaaggttgct 2160
gggggtcatt atgtccaat ggctctcatg aagttggccg cactgacagg tacgtacgtt 2220
tatgaccatc tcacccact gcgggactgg gccacgcgg gcttacgaga ccttgcgtg 2280
gcagttgagc ccgtcgtctt ctctgatatg gagaccaagg ttatcacctg gggggcagac 2340
accgcgcggt gtggggacat catcttgggc ctgcccgtct ccgcccgcag gggggaggag 2400
atacatctgg gacggcgaga cagccttgaa gggcaggggg ggcgactcct cgcgcctatt 2460
acggcctact cccaacagac gcgaggccta cttggtgcga tcatcactag cctcacaggc 2520
cgggacagga accaggtcga gggggaggtc caagtgtctt ccaccgcaac acaatctttc 2580
ctggcgacgt cgtcaatgg cgtgtgttgg actgtctatc atggtgcgg ctcaaagacc 2640
cttgccggcc caaaggggcc aatcacccaa atgtacacca atgtggacca ggacctcgtc 2700
ggctggcaag cgcccccg ggcgcgttcc ttgacaccat gcacctgcgg cagctcggac 2760
ctttacttgg tcacgaggca tgcgatgtc attccggtgc gccggcggg cgacagcagg 2820
gggagcctac tctccccag gcccgctccc tacttgaagg gctcttcggg cggctccactg 2880
ctctgcccct cggggcacgc tgtgggcac tttcgggctg ccgtgtgcac ccgaggggtt 2940
gcgaaggcgg tggactttgt acccgtcgag tctatggaaa ccactatcg gtccccggtc 3000
ttcacggaca actcgtcccc tccggccgta ccgcagacat tccaggtggc ccatctaac 3060
gcccctactg gtagcggcaa gagcactaag gtgcgggctg cgtatgcagc ccaagggtat 3120
aaggtgcttg tctgaaccc gtcctgcgcc gccaccctag gtttcggggc gtatatgtct 3180
aaggcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac caggggtgcc 3240
cccatcagct actccaccta tggcaagttt cttgccgacg gtggttgcctc tgggggcgcc 3300
tatgacatca taatatgtga tgagtgcac tcaactgact cgaccactat cctgggcac 3360
ggcacagtc tggaccaagc ggagacggct ggagcgcgac tcgtcgtgct cgccaccgct 3420
acgctccgg gatcggtcac cgtgccacat ccaaacatcg aggaggtggc tctgtccagc 3480
actggagaaa tcccccttta tggcaaagcc atccccatcg agaccatcaa gggggggagg 3540

```

	cacctcattt	tetgccattc	caagaagaaa	tgtgatgagc	tcgccgcgaa	gctgtccggc	3600
	ctcggactca	atgctgtagc	atattaccgg	ggccttgatg	tatccgtcat	accaactagc	3660
	ggagacgtca	ttgtcgtagc	aacggacgct	ctaatagcgg	gctttaccgg	cgatttcgac	3720
5	tcagtgatcg	actgcaatac	atgtgtcacc	cagacagtcg	acttcagcct	ggacccgacc	3780
	ttcaccattg	agacgacgac	cgtgccacaa	gacgcggtgt	caagctcgca	gcggcgaggc	3840
	aggactggta	ggggcaggat	gggcattttac	aggtttgtga	ctccaggaga	acggccctcg	3900
	ggcatgttcg	attcctcggg	tctgtgcgag	tgetatgacg	cgggctgtgc	ttgggtacgag	3960
	ctcacgcgcg	ccgagacctc	agtttaggtt	cgggcttacc	taaacacacc	agggttgccc	4020
	gtctgccagg	accatctgga	gttctggggg	agcgtcttta	caggcctcac	ccacatagac	4080
	gccattttct	tgtcccagac	taagcaggca	ggagacaact	tcacctacct	ggtagcatat	4140
10	caggctacgg	tgtgcgcagc	ggctcaggct	ccacctccat	cgtgggacca	aatgtggaag	4200
	tgtctcatac	ggctaaagcc	tacgctgcac	gggccaaagc	ccctgctgta	taggctggga	4260
	gccgttcaaa	acgaggttac	taccacacac	cccataacca	aatacatcat	ggcatgcatg	4320
	tcggctgacc	tggaggtcgt	cacgagcacc	tgggtgctgg	taggcggagt	cctagcagct	4380
	ctggccgcgt	attgcctgac	aacaggcagc	gtggtcattg	tgggcaggat	catcttgtcc	4440
	ggaaagccgg	ccatcattcc	cgacagggaa	gtcctttacc	gggagttcga	tgagatggaa	4500
15	gagtgcgcct	cacacctccc	ttacatcgaa	cagggaatgc	agctcgccga	acaattcaaa	4560
	cagaagcгаа	tcgggttgct	gcaaacagcc	accaagcaag	cgagggtcgc	tgctcccgtg	4620
	gtggaattcca	agtggcggac	cctcgaagcc	ttctgggcga	agcatatgtg	gaatttcac	4680
	agcgggatac	aatatttagc	aggcttgctc	actctgcctg	gcaacccgcg	gatagcatca	4740
	ctgatggcat	tcacagcctc	tatcaccagc	ccgctcacca	cccaacatac	cctcctgttt	4800
	aacatcctgg	ggggatgggt	ggccgcccaa	cttgctcctc	ccagcgtcgc	ttctgcttcc	4860
	gtagggcccg	gcacgcgtgg	agcggctgtt	ggcagcatag	gccttgggaa	gggtgctgtg	4920
20	gataattttg	caggttatgg	agcaggggtg	gcaggcgcgc	tcgtggcctt	taaggtcatg	4980
	agcggcgaga	tgcctccac	cgaaggacct	gttaacctac	tccttcttat	cctctccctt	5040
	ggcgccctag	tcgtcggggt	cgtgtgcgca	gcgatactgc	gtcggcacgt	ggggccaggg	5100
	gagggggctg	tgcagtggat	gaaccggctg	atagcgttcg	cttcgcgggg	taaccacgtc	5160
	tcccccacgc	actatgtgcc	tgagagcgac	gctgcagcac	gtgtcactca	gatcctctct	5220
	agtcttacc	tcactcagct	gctgaagagg	cttcaccagt	ggatcaacga	ggactgctcc	5280
25	acgccatgct	ccggctcgtg	gctaagagat	gtttgggatt	ggatatgcac	gggtgtgact	5340
	gatttcaaga	cctggtccca	gtccaagctc	ctgccgcgat	tgcggggagt	ccccttcttc	5400
	tcattgtcaac	gtgggtacaa	gggagctctg	cgggggcagc	gcaltatgca	aaccacctgc	5460
	ccatgtggag	cacagatcac	cggacatgtg	aaaaacggtt	ccatgaggat	cgtggggcct	5520
	aggacctgta	gtaaacctgt	gcattggaaca	ttccccatta	acgcgtacac	cacgggcccc	5580
	tgcacgccct	ccccggcgcc	aaattattct	agggcgctgt	ggcgggtggc	tgctgaggag	5640
	tacgtggagg	ttacgcgggt	gggggatttc	cactacgtga	cgggcacgac	cactgacaa	5700
30	gtaaaagtcc	cgtgtcaggt	tcggccccc	gaattcttca	cagaagtggg	tgggtgtcgg	5760
	ttgcacaggt	acgctccagc	gtgcaaaccc	ctcctacggg	aggaggtcac	attcctgggtc	5820
	gggcttcaat	aatacctggt	tgggtcacag	ctcccatcg	agcccgaaac	ggacgtagca	5880
	gtgtctactt	ccatgctcac	cgacccctcc	cacattacgg	cggagacggc	taagcgtagg	5940
	ctggccaggg	gatctccccc	ctccttggtc	agctcatcag	ctagccagct	gtctgcgcct	6000
	tccttgaagg	caaatgcac	taccgctcat	gactccccgg	acgctgacct	catcgaggcg	6060
35	aacctctgt	ggcggcagga	gatggggggg	aacatcacc	gcgtggagtc	agaaaataag	6120
	gtagttaatt	tggactcttt	cgagccgctc	caagcggagg	aggatgagag	ggaagtatcc	6180
	gttccggcgg	agatcctcgg	gaggtccagg	aaattccttc	gagcgtatgc	catatgggca	6240
	cgcgcggatt	acaacctccc	actgttagag	tcctggaaagg	accgggacta	cgctccctca	6300
	gtggtacacg	ggtgtccatt	gccgcctgcc	aaggccccctc	cgataccacc	tcacggagg	6360
	aagaggacgg	ttgtcctgtc	agaatctacc	gtgtcttctg	ccttggcgga	gctcgccaca	6420
	aagaccttcg	gcagctccga	atcgtcggcc	gtcgacagcg	gcacggcaac	ggcctctcct	6480
40	gaccagccct	ccgacgacgg	cgacgcggga	tcgcagcttg	agtcgtactc	ctccatgccc	6540
	ccccttgagg	gggagccggg	ggatcccgat	ctcagcgacg	ggctcttggtc	taccgttaagc	6600
	gaggaggcta	gtgaggacgt	cgtctgctgc	tcgatgtcct	acacatggac	aggcgccctg	6660
	atcacgccat	gcgctgcgga	ggaaaccaag	ctgcccatac	atgcactgag	caactctttg	6720
	ctccgtcacc	acaacttggt	ctatgctaca	acatctcgca	gcgcaagcct	gcggcagaag	6780
	aaggtcacct	ttgacagact	gcaggctctg	gacgaccact	accgggacgt	gctcaaggag	6840
45	atgaaggcga	aggcgtccac	agttaaggct	aaacttctat	ccgtggaggga	agcctgttaag	6900
	ctgacgcccc	cacattcggc	cagatctaaa	tttggctatg	gggcaaaggga	cgtccggaac	6960
	ctatccagca	aggccgttaa	ccacatccgc	tcctgtgtga	aggacttgct	ggaagacact	7020
	gagacaccaa	ttgacacac	catcatggca	aaaaatgagg	ttttctgcgt	ccaaccagag	7080
	aagggggggc	gcaagccagc	tcgccttacc	gtattcccag	atttgggggt	tcgtgtgtgc	7140
	gagaaaaatg	ccctttacga	tgtggtctcc	acctccctcc	aggccgtgat	gggctcttca	7200
	tacggattcc	aatactctcc	tggacagcgg	gtcagattcc	tgggtgaatgc	ctggaaagcg	7260
50	aagaaatgcc	ctatgggctt	cgcatatgac	accgctgttt	ttgactcaac	ggctcactgag	7320
	aatgacatcc	gtgttgaggga	gtcaatctac	caatgtttgtg	acttggcccc	cgaagccaga	7380
	caggccataa	ggtcgctcac	agagcggctt	tacatcgggg	gccccctgac	taattctaaa	7440
	gggcagaact	gcggctatcg	ccggtgcgcg	gcgagcgttg	tactgacgac	cagctgcggt	7500

EP 1 043 399 A2

aataccctca catgttactt gaaggccgct ggggctgtc gagctgcgaa gctccaggac 7560
 tgcacgatgc tcgtatgcgg agacgacctt gtcgttatct gtgaaagcgc ggggacccaa 7620
 gaggacgagg cgagcctacg ggccttcacg gaggctatga ctagatactc tgccccccct 7680
 5 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgctc ctccaatgtg 7740
 tcagtgcgcg acgatgcacg tggcaaaagg gtgtactatc tcaccogtga cccaccacc 7800
 ccccttgcgc gggctgcgtg ggagacagct agacacactc cagtcaattc ctggctaggc 7860
 aacatcatca tgtatgcgcc caccttggtg gcaaggatga tcctgatgac tcatttcttc 7920
 tccatccttc tagctcagga acaacttgaa aaagccctag attgtcagat ctacggggcc 7980
 10 tgttactcca ttgagccact tgacctacct cagatcattc aacgactcca tggccttagc 8040
 gcattttcac tccatagtta ctctccaggt gagatcaata gggaggcttc atgcctcagg 8100
 aaacttgggg taccgccctt gcgagtcctg agacatcggg ccagaagtgt ccgcgctagg 8160
 ctactgtccc agggggggag ggctgccact tgtggcaagt acctcttcaa ctgggcagta 8220
 aggaccaagc tcaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280
 ttcgttgctg gttacagcgg gggagacata tatcacagcc tgtctcgtgc ccgacccgc 8340
 15 tggttcatgt ggtgcctact cctactttct gtaggggtag gcatctatct actccccaac 8400
 cgatgaacgg ggagctaaac actccaggcc aataggccat cctgtttttt tccctttttt 8460
 ttttctttt ttttttttt ttttttttt ttttttttt ctctttttt tttcctcttt 8520
 ttttctttt ctttctttt gtggctccat cttagcccta gtacgggcta gctgtgaaag 8580
 gtccgtgagc cgcttgactg cagagagtgc tgatactggc ctctctgcag atcaagt 8637
 20
 25
 30
 35
 40
 45
 50
 55

<210> 3
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 3
 gccagccccc gattgggggc gacactccac catagatcac tccccgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtggctcg cggaaaccgt gactacaccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc 240
 10 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgttact gcctgatagg 300
 gtgcttgaga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaaagaaa aaccaaacgt aacaccaacg ggccgcccat gattgaacaa gatggattgc 420
 acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 15 cgtgctggcg cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
 gaagggactg gctgctattg ggcaagtgcc cggggcagga tctcctgtca tctcaccttg 720
 ctccctgccg gaaagtatcc atcatggctg atgcaatgcg gcggtgtcat acgcttgatc 780
 cggtctacct ccatttcgac caccaaagca aacatcgcat cgagcgagca cgtactcgga 840
 tggaaagccg tcttgtcgat caggatgac tggacgaaga gcatcagggg ctgcgcaccg 900
 ccgaactgtt cgcagggctc aaggcgcgca tgcgcgacgg cgaggtatctc gtcgtgaccc 960
 atggcgatgc ctgcttgcgg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 20 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
 ttgctgaaga gcttggcgcc gaatggctg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatcc gcagcgcatc gccttctatc gccttcttgc cgagtctctc tgagtttaaa 1200
 cagaccacaa cggtttccct ctacgggat caattccgca cctctccctc cccccccctc 1260
 aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgcgctcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 25 acgagcattc ctagggtctc ttccctctc gccaaaggaa tgcaaggtct gttgaatgta 1440
 gtgaagggaag cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcgcacctt 1500
 tgcaggcagc ggaaccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg gcacaaacc cagtgccacg ttgtgagttg gatagtgtg 1620
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaa 1680
 gtaccccatc gtatgggata tgatctgggg cctcgggtga catgctttac atgtgtttag 1740
 30 tcgagggtta aaaaagctca gggcccccca accacgggga cgtggtttcc ctttgaaaaa 1800
 cagcataata ccattggccc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcacatca ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagt 1920
 gtctccaccg caacacaatc tttctggcg acctgcgtca atggcggtg ttggactgtc 1980
 tatcatgttg ccggtcaaaa gaccttgcc ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgttg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttccttgaca 2100
 ccattgacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
 35 gtgcgcccgc gggggcagac cagggggagc ctactctccc ccaggcccg tctcctactt 2220
 aagggtctct cgggcggtcc actgctctgc cctcggggc acgctgtggg catctttcgg 2280
 gctgcggtgt gcaccgagg ggttgcgaag gcggtggact ttgtaccctg caggtctatg 2340
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacgcccc actggtagcg gcaagagcac taaggtgccg 2460
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgctcgt cgccgccacc 2520
 40 ctagggttct gggcgatat gtctaaggca catggtatcg accctaaccat cagaaccggg 2580
 gtaaggacca tcaccaggg tgcccccat acgtactcca cctatggcaa gttcttggc 2640
 gacggtggtt gctctgggg cgccatgac atcataatat gtgatgagtg ccaactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcttgacc aagcggagac ggctggagcg 2760
 cgactcgctg tgcctgccac cgtacgcct ccgggatcgg tcacgctgcc acatccaaa 2820
 atcgaggagg tggtctgtgc cagcaactga gaaatccct tttatggcaa agccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 45 gagctcgccg cgaagctgtc cggcctcgga ctcaatgtc tagcatatta ccggggcctt 3000
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaagt 3060
 acgggtctta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtgcgactta gctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
 gtgtcagctc cgcagcggcg aggcaggact ggtaggggca ggtaggcat ttacaggttt 3240
 gtgaactccg gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtctat 3300
 50 gacgcgggct gtgcttggtc cgagctcacg cccgcgaga cctcagttag gttgcgggct 3360
 taactaaaca caccagggtt gcccgctgcg caggaccatc tggagttctg ggagagcgct 3420
 tttacaggcc tcaccacat agacgcccac ttcttgtccc agactaagca ggcaggagac 3480
 aacttcccc acctggtagc ataccaggt acggtgtgcg ccagggtcga ggctccacct 3540

55

	ccatcggtggg	accaaattgtg	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgccccctgc	tgtataggct	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
	atttggggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccgacag	ggaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaag	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgctcc	cgtgggtgaa	tccaagtggc	ggaccctcga	agccttctgg	4020
	gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
10	cctggcaacc	ccgcgatagc	atcactgatg	gcatttcacag	cctctatcac	cagcccgtcc	4140
	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
	ataggcccttg	ggaagggtgc	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
	gcgctcgtgg	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
	ctactccctg	ctatccctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctgcgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatacgc	4500
15	ttcgtctcgc	ggggttaacca	cgtctccccc	acgcactatg	tgcttgagag	cgacgctgca	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
	cagtggatca	acgaggactg	ctccacggca	tgctccggct	cgtggctaag	agatgtttgg	4680
	gattggatat	gcacgggtgtt	gactgatttc	aagacctggc	tcacgtccaa	gctcctgccc	4740
	cgattgcccgc	gagtcctcct	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaaaccac	ctgcccattg	ggagcacaga	tcaccgggaca	tgtgaaaaac	4860
20	ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
	attaacgcgt	acaccacggg	ccctctgcacg	ccctccccgg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	caacgtlaaag	tgcccgtgtc	aggttcgggc	ccccgaatte	5100
	ttcacagaag	tggtgggggt	gcggttgac	aggtacgtct	cagcgtgcaa	acccctccta	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatacaatcc	tggttggtgc	acagctccca	5220
	tgcgagcccg	aaccggacgt	agcagtgtct	acttccatgc	tcaccgaccc	ctccacattt	5280
25	acggcggaga	cggctaagcg	taggtctggc	aggggatctc	ccccctcctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactaccgg	tcagtactcc	5400
	ccggacgctg	acctcatcga	ggcacaectc	ctgtggcggc	aggagatggg	cgggaacatc	5460
	accgcgtggg	agtcaaaaaa	taaggtagta	atlttggaat	ctttcgagcc	gctccaagcg	5520
	gaggaggatg	agagggaagt	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtccctg	5640
30	aaggaccctg	actacgtccc	tccagtggta	cacgggtgtc	cattgccgcc	tgccaaggcc	5700
	cctccgatac	cacctccacg	gaggagaggg	acgggtgtcc	tgtcagaatc	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	cctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtcgt	actcctccat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcagggag	gctagttagg	acgtcgtctg	ctgctcgatg	6000
	tcctacacat	ggacaggcgc	cctgatcagc	ccatgcgtcg	cggaggaaac	caagctgccc	6060
35	atcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcagggt	cctggacgac	6180
	cactaccggg	acgtgtctca	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacattt	cggccagatc	taaattttgg	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
	tggaaggact	tgctggaaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgccct	tatcgtattc	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggg	ctccaccctc	6540
	cctcaggccg	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttccttggtg	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tggtttgact	caacggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgacttgg	cccccgaaagc	cagacaggcc	ataaggctcg	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaatte	taaagggcag	aactgcggct	atcgccggtg	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggtaataacc	ctcacatggt	acttgaaggc	cgtcggggcc	6900
	tgctcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gaggcgagcc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccacac	cagaatacga	cttggagttg	7080
	ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aaggtgttac	7140
	tatctcaccc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
	actccagtcga	attcctggct	aggcaacatc	atcatgtatg	cgcccaactt	gtgggcaagg	7260
50	atgatactga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgttac	tcatttagag	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggttg	cttcatgcct	caggaaaactt	ggggtaccgc	ccttgcgagt	ctggagacat	7500

EP 1 043 399 A2

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
 agcctgtctc gtgcccgacc ccgctgggtc atgtggtgcc tactcctact ttctgtaggg 7740
 gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
 ccatcctggt tttttccctt tttttttttc tttttttttt tttttttttt tttttttttt 7860
 ttttctcctt tttttttcct ctttttttcc ttttctttcc tttggtggct ccatcttagc 7920
 cctagtcacg gctagctgtg aaaggctcgt gagecgcttg actgcagaga gtgctgatac 7980
 10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

EP 1 043 399 A2

<210> 4
<211> 7989
<212> DNA
<213> Hepatitis C Virus

5

10

15

20

25

30

35

40

45

50

55

```

<400> 4
gccagccccc gattgggggc gacactccac catagatcac tccccgtgta ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggtctg cggaaaccgt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaagccc ttgtgtgtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaaa aaccaaagg cgcccatga ttgaacaaga tggattgcac gcaggttctc 420
cgcccgcttg ggtggagagg ctattcgct atgactgggc acaacagaca atcggtgct 480
ctgatgcgcg cgtgttccgg ctgtcagcgc aggggcgccc gggtcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacggcgct tccttgcca gctgtgctcg acgttgtcac tgaagcggga agggactggc 660
tgctattggg cgaagtgcgg gggcaggatc tectgtcac tcaccttgct cctgccgaga 720
aagtatccat catggtgat gcaatggcgc ggctgcatac gcttgatccg gctacctgcc 780
cattcgacca ccaagcgaac catcgcatcg agcagcacg tactcggtat gaagccgggtc 840
ttgtcgatca ggtgatctg gacgaagagc atcaggggct cgcccgacc gaactgttcg 900
ccaggctcaa ggcgcgcatg cccgacggcg aggtctcgt cgtgacccat ggcgatgcct 960
gcttgccgaa tatcatgggt gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1020
tgggtgtggc ggaaccgtat caggacatag cgttggttac ccgtgatatt gctgaagagc 1080
ttggcgcgca atgggctgac cgcttccctg tgctttacgg tatcgccgct cccgattcgc 1140
agcgcatcgc cttctatcgc cttcttgacg agttcttctg agtttaaaaca gaccacaacg 1200
gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc 1260
cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320
ccgtcttttg gcaatgtgag ggcccgaaa cctggccctg tcttcttgac gagcattcct 1380
aggggtcttt cccctctcgc caaaggaaatg caaggtctgt tgaatgtcgt gaaggaaagc 1440
gttccctctg aagcttcttg aagacaaaca acgtctgtag cgacctttg caggcagcgg 1500
aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgttttagtc gaggttaaaa 1740
aacgtctagg ccccccgaac caggggagcg tggttttcct ttgaaaaaca cgataatac 1800
atggcgcta ttacggccta ctcccaacag acgcgaggcc tacttggtcg catcatcact 1860
agcctcacag gccgggacag gaaccaggtc gagggggagg tccaaagtgt ctccaccgca 1920
acacaatctt tcttgggac ctgcgtcaat ggctgtgtgt ggactgtcta tcatggtgcc 1980
ggctcaaaag cccctgcggg cccaaagggc ccaatcaccc aaatgtacac caatgtggac 2040
caggacctcg tcgcttgcca agcgcctccc ggggcggtt ccttgacacc atgcacctgc 2100
ggcagctcgg acctttactt ggtcacagg catgcgatg tcattccggg gcgcggcgcg 2160
ggcgacagca gggggagcct actctccccc aggccgctct cctacttgaa gggctcttcg 2220
ggcggtccac tgctctgccc ctccggggcac gctgtgggca tctttcgggc tgcgtgtgc 2280
acccgagggg ttgcgaaggc ggtggacttt gtaccctcg agtctatgga aaccactatg 2340
cggctccccg tcttcacgga caactcgctc cctccggccg tacccgagac attccagggt 2400
gccatcttac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgatgca 2460
gcccaagggt ataaggtgct tgtctgaac ccgtccgtcg ccgccacct aggtttcggg 2520
gcgtatatgt ctaaggcaca tggatcgac cctaacatca gaaccgggtt aaggaccatc 2580
accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtgttgc 2640
tctggggcg cctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 2700
atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 2760
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 2820
gctctgtcca gcactggaga aatccccctt tatggcaaaag ccatcccat cgagaccatc 2880
aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 2940
aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc 3000
ataccaacta gcggagacgt cattgtcgta gcaacggacg ctctaagac gggctttacc 3060
ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc 3120
ctggaccoga ccttcaccat tgagacgacg accgtgccac aagacgcggt gtcacgctcg 3180
cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga 3240
gaacggccct cgggcatgtt cgattcctcg gttctgtcgc agtgcctatga cgcgggctgt 3300
gcttggtacg agctcacgcc cgccgagacc tcagttaggt tgccggctta cctaaacaca 3360
ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc 3420
accacatag acgcccattt cttgtcccag actaagcagg caggagacaa cttccctac 3480
ctggtagcat accaggctac ggtgtgccc agggctcagg ctccacctcc atcgtgggac 3540

```

caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gccctgtctg 3600
 tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc 3660
 atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3720
 gtcctagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg 3780
 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccgggagttc 3840
 gatgagatgg aagagtgcgc ctccacacctc ccttacatcg aacagggaat gcagctcgcc 3900
 gaacaattca aacagaaggc aatcggttg ctgcaaacag ccaccaagca agcggaggct 3960
 gctgctcccg tggtagaatc caagtggcgg accctcgaa gcttctgggc gaagcatatg 4020
 tggaaattca tcagcgggat acaatattta gcaggctgt ccaactctgc tggcaacccc 4080
 gcgatagcat cactgatggc attcacagcc tctatacca gcccgctcac caccacaatc 4140
 accctcctgt ttaacatcct ggggggatgg gtggccgccc aacttgctcc tcccagcgtc 4200
 gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg 4260
 aaggtgcttg tggatatctt ggcaggttat ggagcagggg tggcaggcgc gctcgtggcc 4320
 ttttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380
 atcctctccc ctggcgccct agtcgtcggg gtcgtgtgcg cagcgatact gcgtcgccac 4440
 gtgggccccg gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgccc 4500
 ggtaaccacg tctccccac gcaactatgt cctgagagcg acgctgcagc acgtgtcact 4560
 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac 4620
 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 4680
 acggtgttga ctgatttcaa gacctggctc cagtccaagc tcctgcccgc attgcccggga 4740
 gtccccctct tctcatgtca acgtgggtac aaggaggtct ggccgggcca cggtcatcatg 4800
 caaaccacct gcccatgtgg agcacagatc accggacatg tgaanaacgg ttccatgagg 4860
 atcgtggggc ctaggacctg tagtaacacg tggcatggaa cattcccatc taacgcgtac 4920
 accacggggc cctgcacgcc ctccccggcg ccaaattatt ctaggggcgt gtggcgggtg 4980
 gctgctgagg agtacgtgga ggttacggcg gtgggggatt tccactacgt gacgggcatg 5040
 accactgaca acgtaaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg 5100
 gatggggtgc ggttgacag gtacgtctca gcgtgcaaac ccctcctacg ggaggaggtc 5160
 acattctctg tcgggctcaa tcaatacttg gttgggtcac agctcccatg cgagcccgaa 5220
 ccggacgtag cagtgtcac tccatgtctc accgacccct ccacattac ggcggagacg 5280
 gctaagcgta ggtggccag gggatctccc ccctccttg ccagctcacc agctagccag 5340
 ctgtctcgcg ctctcctgaa ggcaacatgc actaccgctc atgactcccc ggacgctgac 5400
 ctcatcgagg ccaacctcct gtggcggcag gagatggcg ggaacatcac ccgctggag 5460
 tcagaaaata aggtagtaat tttgactct ttcgagccgc tccaagcgga ggagatgag 5520
 agggaaagtat ccgttccggc ggagatcctg cggaggtcca ggaattccc tcgagcgatg 5580
 ccctatggg cagccccgga ttacaacctt ccactgttag agtccctgaa ggaccggac 5640
 tagctccctc cagtgttaca cgggtgtcca ttgcgcctg ccaaggcccc tccgatacca 5700
 cctccacgga ggaagaggac ggttgcctg tcagaatcta ccgtgtcttc tgccttggcg 5760
 gagctcgcca caaagacctt cggcagctcc gaatcgtcgg ccgtcgacag cggcacggca 5820
 accgctcttc ctgaccgcc ctccgacgac ggcgacggcg gatccgacgt tgagtctga 5880
 tctccatgc cccccctga gggggagccg ggggatcccg atctcagcga cgggtcttgg 5940
 tctaccgtaa gcgaggaggc tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg 6000
 acaggcgccc tgatcacgcc atgcgtgcg gaggaaacca agctgcccac caatgcactg 6060
 agcaactctt tgctccgtca ccacaacttg gtctatgcta caacatctcg cagcgcaagc 6120
 ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tggacgacca ctaccgggac 6180
 gtgctcaagg agatgaaggc gaaggcgctc acagttaagg ctaaaactct atccgtggag 6240
 gaagcctgta agctgacgac ccacattcg gccagatcta aatttgcta tggggcaaa 6300
 gacgtccgga acctatccag caaggcgtt aaccacatcc gctccgtgtg gaaggactg 6360
 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 6420
 gtccaaccag agaagggggg ccgcaagcca gctcgcctta tctattccc agatttgggg 6480
 gttcgtgtgt gcgagaaaat ggcctcttac gatgtggtc ccccccctcc tcaggccgtg 6540
 atgggctctt catacggatt ccaatactct cctggacagc gggctcgagt cctggtgaat 6600
 gcctggaag cgaagaaatg cctatgggc ttgcctatg acaccgctg ttttgactca 6660
 agcgtcactg agaatgacat ccgtgttgag gactcaatct accaatgttg tgacttggcc 6720
 ccggaagcca gacaggccat aaggtcgtc acagagcggc tttacatcgg gggccccctg 6780
 actaattcta aagggcagaa ctgcggctat cgcgggtgcc gcgcgagcgg tgtactgacg 6840
 accagctgag gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgag 6900
 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 6960
 gcggggaccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac 7020
 tctgcccccc ctggggaccc gcccaacca gaatacgaact tggagtgtat aacatcatgc 7080
 tctccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgacta tctcaccgtg 7140
 gaccccccca ccccccttgc gcgggctgag tgggagacag ctagacacac tccagtcaat 7200
 tcttgctag gcaacatcat catgtatgag ccacacttgt gggcaaggat gatcctgatg 7260
 actcatttct tctccatcct cttagctcag gaacaacttg aaaaagccct agattgtcag 7320
 atctacgggg cctgttactc cattgagcca cttgacctac ctcatgcat tcaacgactc 7380
 catggcctta gcgcatttct actccatagt tactctccag gtgagatcaa taggttggct 7440
 tcatgcctca ggaacttgg ggtaccgccc ttgcgagttc ggagacatcg ggcagaagt 7500

EP 1 043 399 A2

```

gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 7560
aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 7620
ttatccagct ggttcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt 7680
5  gcccgacccc gctggttcat gtggtgccta ctctacttt ctgtaggggt aggcattctat 7740
ctactcccca accgatgaac ggggagctaa aactccagg ccaataggcc atcctgtttt 7800
tttccctttt ttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt 7860
tttttcctct ttttttctt ttctttcctt tggtaggtcc atcttagccc tagtcacggc 7920
tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 7980
10 agatcaagt

```

15

20

25

30

35

40

45

50

55

EP 1 043 399 A2

<210> 5
 <211> 8649
 <212> DNA
 <213> Hepatitis C Virus

5

```

<400> 5
gccagccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccagta gtgttggtgc gcgaaaggcc ttgtgtgact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaaa aaccaaacgt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttcccttgc cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgcgca gaaagtatcc atcatggctg atgcaatgcy gcggctgcat acgcttgatc 780
cggctacctg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgctgat caggatgatc tggacgaaga gcatcagggg ctgcgcgacg 900
ccgaactgtt cgcagagctc aagcgcgcca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttgccggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctccccgatt gcagcgatc gccctctatc gcccttctga cgagtctctc tgagttaaaa 1200
cagaccacaa cgttttccct ctacgggatc caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
tccaccatat tgcggtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcatte ctagggtctt ttccctctcc gccaaaggaa tgcaaggctc gttgaatgtc 1440
gtgaaggaag cagtctctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatatac ctgcaaaagg gcacaaacc cagtgccacg ttgtgagttg gatagtgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcacagaag 1680
gtacccattt gtatgggacg tgatctgggg cctcgggtga catgctttac atgtgtttag 1740
tcgagggttaa aaaacgtcta ggcgcccgca accacgggga cgtggtttcc ctttgaaaaa 1800
acgataata ccatggaccg ggagatggca gcacgtgcg gagcgcggtt ttctgtaggt 1860
ctgatactct tgacctgtgc accgcactat aagctgttcc tcgctaggct catatgggtg 1920
ttacaatat ttatcacccg gcccgaggca cacttgcaag tgtggatccc ccccccaac 1980
gttcgggggg gcccgatctc cgtcatctcc ctacgtgcg cgatccaccc agagctaata 2040
tttaccatca ccaaaatctt gctcgccata ctcggtccac tcatggtgct ccaggctggt 2100
ataaccaaag tgcgctactt cgtgcgcgca cacgggctca ttcgtgcacg catgctgggt 2160
cggaaggttg ctgggggtca ttatgtccaa atggctctca tgaagtggc cgactgaca 2220
ggtacgtacg tttatgacca tctcacccca ctgcgggact ggcccccacg ggccctacga 2280
gaccttgccg tggcagttga gcccgctgct ttctctgata tggagaccaa ggttatcacc 2340
tggggggcag acaccgcggc gtgtggggac atcatcttg gccctgccgt ctccgcccgc 2400
agggggaggg agatacatct gggaccggca gacagccttg aagggcaggg gtggcgactc 2460
ctcgcgccta ttacggccta ctcccaacag acgcgaggcc tacttgctg catcatcact 2520
agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtgtt ctccaccgca 2580
acacaatctt tccctggcgac ctgcgtcaat ggcgtgtgtt ggaactgtct tcatggtgac 2640
ggctcaaaga ccccttgccg cccaaagggc ccaatcaccc aaatgtacac caatgtggac 2700
caggacctcg tcggctggca agcgcgcccc ggggcgcggt ccttgacacc atgcacctgc 2760
ggcagctcgg accttactt ggtcacgagg catgcccgtg tcatctcggg gcgcgcggcg 2820
ggcgacagca gggggagcct actctcccc aggcgcgtct cctacttgaa gggctcttcg 2880
ggcgggtccac tgcctctgcc ctcggggcac gctgtgggca tctttcgggc tgcgctgtgc 2940
acccgagggg ttgcgaaggc ggtggacttt gtaccctgc agtctatgga aaccactatg 3000
cgttcccccg tcttcacgga caactcgtcc cctccggccg taccgcagac attccaggtg 3060
gcccattctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 3120
gcccaagggt ataagggtct tgtcctgaac ccgtccgtcg ccgccacct aggtttcggg 3180
gcgtatatgt ctaaggcaca tggtatcgac cctaacatca gaaccgggtt aaggaccatc 3240
accacgggtg cccccatcac gtactccacc tatggcaagt ttcttgccga cgttggttgc 3300
tctgggggcy cctatgacat cataatatgt gatgagtgc actcaactga ctcgaccact 3360
atcctggggc tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg 3420
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggagggt 3480
gctctgtcca gcaactggaga aatccccctt tatggcaaa ccatcccatc cgagaccatc 3540
    
```

55

	aaggggggga	ggcacctcat	tttctgccat	tccaagaaga	aatgtgatga	gctcgccgcg	3600
	aagctgtccg	gctcgggact	caatgtctga	gcataattacc	ggggccttga	tgtatccgtc	3660
	ataccaacta	gcggagacgt	cattgtcgta	gcaacggacg	ctctaattgac	gggctttacc	3720
5	ggcgatttcg	actcagtgat	cgactgcaat	acatgtgtca	cccagacagt	cgacttcacg	3780
	ctggaccoga	ccttcaccat	tgagacgacg	accgtgcac	aagacggcgt	gtcacgctcg	3840
	cagcggcgag	gcaggactgg	taggggcagg	atgggcattt	acaggtttgt	gactccagga	3900
	gaacggccct	cgggcatggt	cgattcctcg	gttctgtgcg	agtgtctatga	cgcgggctgt	3960
	gcttggtagc	agctcacgcc	cgcgagaccc	tcagttaggt	tgcgggctta	cctaaacaca	4020
	ccaggggtgc	ccgtctgcca	ggaccatctg	gagttctggg	agagcgtctt	tacaggcctc	4080
	acccacatag	acgcccattt	ctgtgccag	actaagcagg	caggagacaa	cttcccctac	4140
10	ctggtagcat	accaggctac	ggtgtgcgcc	agggtcagg	ctccacctcc	atcgtgggac	4200
	caaagtgtga	agtgtctcat	acggctaag	cctacgctgc	acggggcaac	gcccctgctg	4260
	tataggctgg	gagccgttca	aaacgaggtt	actaccacac	accccataac	caaataacac	4320
	atggcatgca	tgtcggctga	cctggaggtc	gtcacgagca	cctgggtgct	ggtaggcgga	4380
	gtcctagcag	ctctggccgc	gtattgctcg	acaacaggca	gcgtgggtcat	tgtgggcagg	4440
	atcatcttct	ccgaaaagcc	ggccatcatt	cccagacagg	aagtcttcta	ccgggagttc	4500
15	gatgagatgg	aagagtgcgc	ctcacacctc	ccttacatcg	aacagggaat	gcagctcgcc	4560
	gaacaatttc	aacagaaggc	aatcgggttg	ctgcaaacag	ccaccaagca	agcggaggct	4620
	gctgtctccc	tgttgggaat	caagtggcgg	accctcgaag	ccttctgggc	gaagcatatg	4680
	tggattttca	tcagcgggat	acaatattta	gcaggcttgt	ccactctgcc	tggcaacccc	4740
	gcgatagcat	cactgatggc	attcacagcc	tctatacca	gcccgctcac	caccaacat	4800
	acccctctgt	ttaacatcct	ggggggatgg	gtggccgccc	aacttgcctc	tcccagcgct	4860
	gcttctgtct	tcgtaggcgc	cgcatcgct	ggagcggctg	ttggcagcat	aggccttggg	4920
20	aaggtgcttg	tggatatttt	ggcaggttat	ggagcagggg	tggcagcgcg	gctcgtggcc	4980
	tttaaggtca	tgagcggcga	gatgccctcc	accgagagac	tggttaaact	actccctgct	5040
	atcctctccc	ctggcgccct	agtctgcggg	gtcgtgtgcg	cagcgatact	gcgtcggcac	5100
	gtgggcccag	gggagggggc	tgtgcagtgg	atgaaccggc	tgatagcgtt	cgcttcggcg	5160
	ggtaaccacg	tctccccac	gcactatgtg	cctgagagcg	acgctgcagc	acgtgtcact	5220
	cagatctctc	ctagtcttac	catcactcag	ctgctgaaga	ggcttcacca	gtggatcaac	5280
25	gaggactgct	ccaacgcatg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	5340
	acgggtgttg	ctgatttcaa	gacctggctc	cagtcacaagc	tcctgcgcgc	attgccggga	5400
	gtccccctct	tctcatgtca	acgtgggtac	aagggagtct	ggcggggcga	cggcatcatg	5460
	caaaacacct	gcccattgtg	agcacagatc	accggacatg	tgaaaaacgg	ttccatgagg	5520
	atcgtggggc	ctaggacctg	tagtaaacag	tggcatggaa	cattccccat	taacgcgtac	5580
	accacggggc	cctgcacgcc	ctccccggcg	ccaaattatt	ctagggcgct	gtggcgggtg	5640
	gctgctgagg	agtacgtgga	ggttacgcgg	gtgggggatt	tcactacagt	gacgggcagt	5700
30	accactgaca	acgtaaagtg	cccgtgtcag	gttcgggccc	ccgaattctt	cacagaagtg	5760
	gatgggggtg	gggtgcacag	gtacgctcca	gcgtgcaaac	ccctcctacg	ggaggaggtc	5820
	acattctctg	tcgggctcaa	tcaataacctg	gttgggtcac	agctcccatg	cgagcccgaa	5880
	ccggacgtag	cagtgtctac	ttccatgctc	accgacccct	cccacattac	ggcgagacg	5940
	gctaagcgta	ggctggccag	gggatctccc	ccctccttgg	ccagctcatc	agctagccag	6000
	ctgtctgcgc	cttctctgaa	ggcaacatgc	actaccgctc	atgactcccc	ggacgctgac	6060
35	ctcatcgagg	ccaacctcct	gtggcggcag	gagatgggcg	ggaacatcac	ccgcgtggag	6120
	tcagaaaaata	aggtagtaat	tttggactct	ttcgagccgc	tcgaagcgga	ggaggatgag	6180
	agggaagtat	ccgttccggc	ggagatcctg	cggagggtcca	ggaaattccc	tcgagcgatg	6240
	cccatatggg	cacgcccggg	ttacaacctc	ccactgttag	agtcctggaa	ggacccggac	6300
	tacgtccctc	cagtgtgtaca	cgggtgtcca	ttgcgcgctg	ccaaggcccc	tccgatacca	6360
	cctccacgga	ggaagaggac	ggttgtctcg	tcagaatcta	ccgtgtcttc	tgccttggcg	6420
	gagctcgcca	caaagacctt	cggcagctcc	gaatcgctcg	ccgtcgacag	cggcacggca	6480
40	acggcctctc	ctgaccagcc	ctccgacgac	ggcgacgcgg	gatccgacgt	tgagtctgtc	6540
	tctcccatgc	cccccttga	gggggagccg	ggggatcccg	atctcagcga	cgggtcttgg	6600
	tctacccgtg	gcgaggaggg	tagtgaggac	gtcgtctgct	gctcgtatgc	ctacacatgg	6660
	acagcgcccc	tgatcacgcc	atgcgctcg	gaggaaaacca	agctgcccat	caatgcactg	6720
	agcaactctt	tgtcccgcca	ccacaacttg	gtctatgcta	caacatctcg	cagcgcaagc	6780
	ctgcggcaga	agaaggtcac	ctttgacaga	ctgcaggctc	tggacgacca	ctaccgggac	6840
45	gtgctcaagg	agatgaaggc	gaaggcgtcc	acagttaagg	ctaaacttct	atccgtggag	6900
	gaagcctgta	agctgacgcc	cccacattcg	gccagatcta	aatttggcta	tggggcaag	6960
	gacgtccgga	acctatccag	caaggccgtt	aaccacatcc	gctccgtgtg	gaaggacttg	7020
	ctggaagaca	ctgagacacc	aattgacacc	accatcatgg	caaaaaatga	ggttttctgc	7080
	gtccaaccag	agaagggggg	ccgcaagcca	gctcgcttta	tcgtattccc	agatttgggg	7140
	gttcgtgtgt	gcgagaaaat	ggccctttac	gatgtggtct	ccaccctccc	tcaggccgtg	7200
	atgggctctt	catacggatt	ccaatactct	cctggacagc	gggtcgagtt	cctggtgaat	7260
50	gcttggaag	cgaagaaatg	ccctatgggc	ttcgcatatg	acacccgctg	ttttgactca	7320
	acggtcactg	agaatgacat	ccgtgttgag	gagtcfaatc	accaatgttg	tgacttggcc	7380
	ccgaagcca	gacaggccat	aaggtcgctc	acagagcggc	tttacatcgg	gggccccctg	7440
	actaatctta	aagggcagaa	ctgcggctat	cgcgggtgcc	gcgcgagcgg	tgtactgacg	7500

EP 1 043 399 A2

	accagctg	gtaataccct	cacatgttac	ttgaaggccg	ctgcggcctg	tcgagctg	7560
	aagctccagg	actgcacgat	gctcgtatgc	ggagacgacc	ttgtcgttat	ctgtgaaagc	7620
	gcggggaccc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7680
5	tctgcccccc	ctggggaccc	gccccaaacca	gaatacgact	tggagttgat	aacatcatgc	7740
	tctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcaccctg	7800
	gaccccaacca	ccccccctgc	gcgggctg	tgaggagacag	ctagacacac	tccagtcaat	7860
	tcttgctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7920
	actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	7980
10	atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	8040
	catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	taggggtgct	8100
	tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	8160
	gtccgcgcta	ggctactgtc	ccaggggggg	agggctgcca	cttgtggcaa	gtacctcttc	8220
	aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgct	ccagttggat	8280
	ttatccagct	ggttcgttgc	tggttacagc	gggggagaca	tatatcacag	cctgtctcgt	8340
15	gcccgaaccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	8400
	ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	8460
	tttccctttt	tttttttctt	tttttttttt	tttttttttt	tttttttttt	ttctcctttt	8520
	tttttctctt	ttttttctct	ttctttctct	tggtggctcc	atcttagccc	tagtcacggc	8580
	tagctgtgaa	aggtccgtga	gccgcttgac	tgagagaggt	gctgatactg	gcctctctgc	8640
20	agatcaagt						8649
25							
30							
35							
40							
45							
50							
55							

<210> 6
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 6
 gccagccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtggtctg cggaaccggg gactacaccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgtgact gcctgatagg 300
 gtgcttgcca gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaacc 360
 ctcaaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccttga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcagcggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctccctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggagccggg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcagag 900
 ccgaactggt cgcagagctc aaggcgcgca tgcccgcagg cgaggatctc gtcgtgacct 960
 atggcgatgc ctgcttgccc aatatcatgg tggaaaatgg ccgcttttct gattcatcgc 1020
 actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
 ttgctgaaga gcttggcggc gaattgggtg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200
 cagaccacaa cggtttccct ctacggggat caattccgca cctctccctc cccccccct 1260
 aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgcggtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggtctc ttccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
 gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgacctt 1500
 tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg ggcacaaacc cagtgccacg ttgtgagttg gatagttgtg 1620
 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tcccagaag 1680
 gtacccattt gtatgggacg tgatctgggg cctcgggtga catgctttac atgtgtttag 1740
 tcgagggttaa aaaacgtcta gggcccccga accacgggga cgtgggtttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcatcatca ctagcctcac agggccggac aggaaccagg tcgaggggga ggtccaaagt 1920
 gttccaccgg caacacaatc ttctctggcg acctgcgtca atggcggtgt ttggactgtc 1980
 tatcatggtg ccggtcaaaa gaccttgcg ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttcttgaca 2100
 ccatgcacct gcgagcagct ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
 gtgcgcgggc ggggcgacag cagggggagc ctactctccc ccaggcccg ctctacttg 2220
 aagggctctt cggcggttgc actgctctgc cctcggggc atgctgtggg catctttcgg 2280
 gctgcgctgt gcaaccgagg ggttgcgaag gcggtggact ttgtacccgt cgagtctatg 2340
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acatttccagg tggcccatct acacgcccct actggttagcg gcaagagcac taaggtgccg 2460
 gctgcgtatg cagcccaagg gtataaggty cttgtcctga acccgtccgt cgcgcacc 2520
 ctagggttct ggcggtatat gtctaaggca catggtatcg accctaaccat cagaaccggg 2580
 gtagggacca tcaccacggg tgcccccac acgtactcca cctatggcaa gttcttgcc 2640
 gacggtgggt gctctggggg cgcctatgac atcataatat gtgatgagt ccactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcgagac ggctggagcg 2760
 cgactcgtcg tgcctgcac cgtacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct ttataggcaa agccatcccc 2880
 atcgcgacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctcggcg cgaagctatc cggcctcgga ctcaatgtg tagcatatta ccgggcccct 3000
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgtctaatg 3060
 acgggcttta ccggcgattt cgaactcagt atcgactgca atacatgtgt caccagaca 3120
 gtcgacttca gccctggacc gaccttcacc attgagacga cgaccgtgcc acaagacggc 3180
 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
 gtgactccag gagaacggcc ctccggcattg ttcgattcct cggttctgtg cgagtgtat 3300
 gacgcgggct gtgcttggtg cgaactcagc ccgcgcgaga cctcagttag gttgcgggct 3360
 tacctaataa caccagggct gcccgctctg caggaccatc tggagttctg ggagagcgtc 3420
 ttacagggcc tcacccacat agacgcccac ttcttgtccc agactaggca ggcaggagac 3480
 aacttccccct acctggtagc ataccaggct acggtgtgcg ccagggctca ggctccacct 3540

55

EP 1 043 399 A2

	ccatcgtagg	accaaagtgt	ggagtggtct	atacggttaa	agcctacgct	gcacgggcca	3600
	acgccccctg	tgtataggct	gggagccggt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtgggt	3780
	attgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccgcagc	ggaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtg	gcctcacacc	tcctttacat	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaag	gcaatcggtg	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgctcc	cgtgtgggaa	tccaagtggc	ggaccctcga	agccttctgg	4020
	gcgaagcata	tgtggaattt	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
	cctggcaacc	ccgcgatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgctc	4140
10	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctcccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
	atagcccttg	ggaaggtgct	tgtgatat	ttggcagggt	atggagcagg	gggtggcagg	4320
	gcgctcgtag	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctgggttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgagcgata	4440
	ctgcgtcgcc	acgtgggccc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
15	ttcgcttcgc	ggggtaacca	cgtctccccc	acgcactatg	tgtctgagag	cgacgctgca	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggtctcac	4620
	cagtggatca	acgaggaact	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
	gattggatat	gcacgggtgt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccc	4740
	cgattgcggg	gagtcctcct	cttctcatgt	caacgtgggt	acaaggaggt	ctggcggggc	4800
	gacggcatca	tgcaaacca	ctgcccattg	ggagcacaga	tcaccggaca	tgtgaaaaac	4860
	ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
20	attaacgcgt	acaccacggg	ccctcgacag	ccctccccgg	cgccaaatta	ttctagggcg	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	caacgtaaa	tgcccggtgc	aggttcgggc	ccccgaattc	5100
	ttcacagaag	tggtatgggt	gcggttgac	aggtacgctc	cagcgtgcaa	acccctccta	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctccca	5220
	tgccggcccg	aaccggacgt	agcagtgtgc	acttccatgc	tcaccgaccc	ctccacattc	5280
25	acggcggaga	cggtcaagcg	taggtctggc	aggggatctc	ccccctcctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccttg	aaggcaacat	gcactaccgc	tcagtactcc	5400
	ccggacgctg	acctcatcga	ggccaaacct	ctgtggcggc	aggagatggg	cgggaaacat	5460
	acccgctggg	agtcagaaaa	taaggtagta	alittggact	ctttcgagcc	gctccaagcg	5520
	gaggaggatg	agagggaagt	atccgttccg	gcggagatcc	tgccgaggtc	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgcccg	gattacaacc	ctccactggt	agagtctctg	5640
30	aaggaccggg	actacgtccc	tccagtggta	cacgggtgtc	cattgcgcgc	tgccaaagcg	5700
	cctccgatcc	cacctccacg	gagggagagg	acggttgtcc	tgtcagaatc	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacaaagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtcgt	actcctccat	gccccccctt	gagggggagc	cggggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gctagttagg	acgtcgtctg	ctgctcgatg	6000
35	tcctacacat	ggacaggcgc	cctgatcacg	ccatgcgctg	cggaggaaac	caagctgccc	6060
	gtcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
	caactaccggg	acgtgtctca	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacattt	cggccagatc	taaatattgg	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	ccgctccgtg	6360
	tggaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
40	gaggttttct	gcgtcccaacc	agagaagggg	ggccgcgaag	cagctcgcc	tatcgtattc	6480
	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
	cctcaggccg	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctgttga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacaccgcg	6660
	tgttttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgacttgg	cccccgaaag	cagacaggcc	ataaggctcg	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaatcc	taaagggcag	aactgcggct	atcgcgggtg	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggttaatacc	ctcacatgtt	atttgaaggc	cgttcgggcc	6900
	tgctcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gagggcgagc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccaaac	cagaatacga	cttggagttg	7080
	ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aaggggtgtac	7140
	tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
	actccagtca	attcctggct	aggcaacatc	atcatgtatg	cgccacacct	gtgggcaagg	7260
50	atgatcctga	tgaactattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgttac	tcatttgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagegcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgcct	caggaaactt	ggggtaaccg	ccttgcgagt	ctggagacat	7500

55

EP 1 043 399 A2

```

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
5 agcctgtctc gtgcccgcgc ccgctgggtc atgtgggtgc tactcctact ttctgtaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctggt tttttccctt tttttttttc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttcct ctttttttcc ttttctttcc tttgggtggc ccatcttagc 7920
cctagtcacg gctagctgtg aaaggctcgt gagecgcttg actgcagaga gtgctgatac 7980
10 tggcctctct gcagatcaag t 8001

```

15

20

25

30

35

40

45

50

55

<210> 7
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 7
 gccagccccc gattgggggc gacactccac catagatcac tccccgtgta ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag 180
 gacgacgggg tcctttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgtgact gcctgatagg 300
 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaagaaa aaccaaacgt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggccc ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgtttgc actgaagcgg 660
 gaagggactg gctgctattg ggccgaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctccctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
 cggctacactg ccatttcgac caccaaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggaaagccg tcttgtcgat caggatgatc tggacgaaga gcacagggg ctgcgcaccg 900
 ccgaactgtt cgcagggctc aagcgcgcca tgcgcgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgcttgcg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
 ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatcc gcagcgatc gcccttctat gcccttctga cgagtctctc tgagtttaaa 1200
 cagaccacaa cgttttccct ctacgggatc caattccgcc cctctccctc cccccccctc 1260
 aacgttactg gccgaagccg cttggaataa ggccgggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcctgtctt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggtctc ttccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
 gtgaaggaag cagttctctc ggaagctctc tgaagacaaa caacgtctgt agcgaccctt 1500
 tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg gcgacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
 gaagagtgca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgccagaag 1680
 gtacccattt gtatgggata tgatctgggg cctcggtgca catgttttac atgtgtttag 1740
 tcgagggtta aaaacgtcta gggcccccca accacgggga cgtggtttc ctttgaaaaa 1800
 cagcataata ccatgggaca gaatcctaaa cctcaaaaga aaaccaaacy taacaccaac 1860
 cgcgcgccac aggcagctca gttcccgggc ggtgggtcaga tcgtcgggtg agtttacctg 1920
 ttgcgcgcga ggggccccag gttgggtgtg cgcgcgacta ggaagacttc ctagcggctg 1980
 caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
 gctcagcccc ggtacccctg gccctctat ggcaatgagg gcttgggggt ggacaggtgg 2100
 ctccgtgcac cccgtgtgctc tcggcctagt tggggcccca cgacccccg gcgtaggtcg 2160
 cgcaatttgg gtaaggtcat cgataccctc acgtgcggct tcgcccagct catgggggtac 2220
 attccgctcg tcggcgcccc cctagggggc gctgccaggg cctcgcgcca tggcgctcgg 2280
 gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340
 ttcccttttg ctttctgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
 gtatccggag tgtacctgtt cagcaacgac tgctccaacg caagcattgt gtatgaggca 2460
 gcggacatga tcatgcatac ccccgggtgc gtgccctgcg ttccgggaga caactcctcc 2520
 cgtgtctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
 acgatacgac gccatgtcga tttgctcgtt gggcgggctg ctctctgctc cgctatgtac 2640
 gtgggagatc tctcgcgatc tgttttctc gtcgcccagc tgttcaacct ctgcctcgc 2700
 cggcacgaga cagtacagga ctgcaattgc tcaatatac ccggccacgt gacaggtcac 2760
 cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
 ttactccgga tcccacaagc tgtcgtggat atggtggcgg ggccccattg gggagtcta 2880
 gcgggccttg cctactatcc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
 ctcttttgcg gctgtgacgg gggaaacctat gtgacagggg ggacgatgac caaaaacacc 3000
 ctcgggatta cgtccctctt ttcaaccggg tcatcccaga aaatccagct tgtaaaacacc 3060
 aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
 ttcccttctg cgtgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
 gccagctgca gcccacatga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
 tcacacagct cggaccagag gccttattgt tggcaactac caccocggcc gtgcggtatc 3300
 gtaccocggc cgcaggtgtg tggctccagt tactgcttca ccccaagccc tgtcgtgggt 3360
 gggacgacgg accggttcgg cgtccctacg tacagtgtgg gggagaatga gacggacgtg 3420
 ctgctcttta acaacacgag ccgcgcgcaa ggcaactggt ttggctgtac atggatgaat 3480
 agcactgggt tcaccaagac gtgcgggggc ccccgctgta acatcggggg gatcggcaat 3540

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
 tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actaccata caggctttgg 3660
 cactaccctt gcactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
 gagcacaggg tcgaagccgc atgcaattgg actcaggag agcgtttaa cctggaggac 3780
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtgga ggtattgccc 3840
 tgttcttcca ccacctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatacctgta cggatagagg tcggcgggtt tctcttttgc aatcaaatgg 3960
 gagtatgtcc tgttgetctt ccttcttctg gcggacgcgc gcgtctgtgc ctgcttgtgg 4020
 atgatgctgc tgatagctca agctgaggcc gccctagaga acctgggtgt cctcaacgcg 4080
 gcatccgtgg ccggggcgca tggcattctc tcttctctcg tgttcttctg tctgctctgg 4140
 tacatcaagg gcaggctggt ccttggggcg gcatatgccc tctacggcgt atggccgcta 4200
 ctctctgctcc tgcgtggcgtt accaccacga gcatacggca tggaccggga gatggcagca 4260
 tcgtgcggag gcgcgggttt cgtaggctctg atactcttga ccttgtcacc gcaactataag 4320
 ctgttctctcg ctaggctcat atgggtggtta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggtaccccc cctcaacgtt cggggggggc gcgatgcctg catcctctc 4440
 acgtgcgcga tccaccaga gctaattctt accatcacca aaatcttctg cgcatactc 4500
 ggtccactca tgggtctcca ggctggtata accaaagtgc cgtacttctg gcgcgcacac 4560
 gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620
 gctctcatga agttggccgc actgacaggt acgtacgttt atgacctct caccctactg 4680
 cgggactggg cccacgcggg cctacgagac cttgcgggtg cagttgagcc cgtcgtcttc 4740
 tctgatattg agaccaaggt tatcacctgg ggggcagaca ccgcgcgcgt tggggacatc 4800
 atcttggggc tgcccgtctc cgcgcgcagg gggaggagga tacatctggg accggcagac 4860
 agccttgaag ggcaggggtg gcgactctct cgcctatta cggcctaact ccaacagacg 4920
 cgaggcctac ttggctgcat catcactagc ctacacggcc gggacaggaa ccaggctcag 4980
 ggggaggtcc aagtggcttc caccgcaaca caatcttctc tggcgacctg cgtcaatggc 5040
 gtgtgttggc ctgtctatca tgggtgcggc tcaaagacct tggccggccc aaagggccca 5100
 atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
 gcgcgttctc tgacaccatg cacctgcggc agctcggacc tttacttggg cacgaggcat 5220
 gccgatgtca ttccggtgcg ccggcggggc gacagcaggg ggagcctaact ctccccagg 5280
 cccgtctctc acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcatgct 5340
 gtgggcatct ttccggctgc cgtgtgcacc cgagggttg cgaaggcggg ggaacttcta 5400
 cccgtcgagt ctatggaaac cactatgcgg tccccggtct tcacggacaa ctgctcccct 5460
 ccggccgtac cgcagacatt ccaggtggcc catctacacg cccctaactg tagcggcaag 5520
 agcactaagg tgccggctgc gtatgcagcc caagggtata aggtgcttct cctgaacccg 5580
 tccgtcgcgg ccacctagg tttcggggcg tatatgtcta aggcacatgg tatcgacct 5640
 aacatcagaa ccggggtagg gacctacacc acgggtgccc ccatcacgta ctccacctat 5700
 ggcaagtctc ttgccgacgg tgggtgctct gggggcgcct atgacatcat aatatgtgat 5760
 gagtgcacct caactgactc gacctatctc ctgggcatcg gcacagctct ggaccaagcg 5820
 gagacggctc gagcgcgact cgtctgtctc gccaccgcta cgcctccggg atcggtcacc 5880
 gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
 ggcaaaagcca tcccatcgc gacctcaag ggggggaggc acctcatttt ctgccattcc 6000
 aagaagaaat gtgatgact cgcgcgaaag ctatccggcc tcggactcaa tgcgttagca 6060
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgcgttagca 6120
 acggacgctc taatgacggg ctttaccggc gatctcgact cagtatcga ctgcaataca 6180
 tgtgtcacc agacagtcga cttcagctcg gacctgacct tcaccattga gacgacgacc 6240
 gtgcacaaag acgcggtgtc acgctcgca cggcgaggca ggaactgtag gggcaggatg 6300
 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttoga tctctcgggt 6360
 ctgtgcgagt gctatgacgc gggctgtgct tggtagcagc tcacgcccgc cgagacctca 6420
 gttaggttga gggcttacct aaacacacca gggctgcccg tctgccagga ccatctggag 6480
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtccagact 6540
 aggcaggcag gagacaaatt cccctaactg gtagcatacc aggcatacgt gtgcgccagg 6600
 gctcaggctc cactccatc gtgggaccaa atgtgggagt gtctcatac gctaaagcct 6660
 acgctgcacg ggccaacgcc cctgctgtat aggcgggag ccgttcaaaa cgaggttact 6720
 accacacacc ccataacca atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
 acgagcactt ggggtctggt aggcggagtc ctagcagctc tggccgcgta ttgcttgaca 6840
 acaggcagcg tggctattgt gggcaggatc atcttgtccg gaaagccggc catcattccc 6900
 gacagggaag tctttaccg ggagtctgat gagatggaag agtgccctc acacctccct 6960
 tacatcgaa agggaaatga gctcgcga caattcaaac agaaggcaat cgggttgctg 7020
 caaacagcaa ccaagcaagc ggaggtctgt gctcccggtg tggaaatcaa gtggcggacc 7080
 ctcaagacct tctgggcgaa gcatatgtgg aatttcatca gcgggataca atatttagca 7140
 ggcttgtcca ctctgcctgg caaccccgcg atagcatcac ctctgttata acatcctggg gggatgggtg 7200
 atcaccagcc cgtccaccac ccaacatacc cagcgtgct tctgctttcg taggcgcgg catcgctgga 7260
 gcccccacac ttgtctctcc ccttgggaag gtgcttgtgg atattttggc aggttatgga 7320
 gcggctgttg gcagcatagg ccttgggaag aaggtcatga gcggcgagat gccctccacc 7380
 gcagggttgg caggcgcgct cgtggccttt ccttcccctg gcgcctagt cgtcggggtc 7440
 gaggacctgg ttaacctact ccttgcctatc ctctcccctg gcgcctagt cgtcggggtc 7500

	gtgtgcgcag	cgatactgcg	tccgcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgtttcg	ttcgcggggt	aaccacgtct	ccccacgca	ctatgtgtct	7620
	gagagcgacg	ctgcagcagc	tgtaactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagaggc	ttcaccagtg	gatcaacgag	gactgctcca	cgccatgctc	cggtctgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaacat	tccccattaa	cgctacacc	acggggccct	gcacgcctc	cccgcgccca	8040
10	aattattcta	gggcgctgtg	gcgggtgggt	gctgaggagt	acgtggaggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaagtcccc	gtgtcagggt	8160
	ccggcccccg	aattcttcac	agaagtggat	gggtgctggg	tgcaacagta	cgctccagcg	8220
	tgcaaacccc	tcctacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctgggt	8280
	gggtcacagc	tcccatgctg	gccgaacccg	gacgtagcag	tgctcacttc	catgctcacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggg	tggccagggg	atctccccc	8400
15	tccctggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	accctcatg	actccccgga	cgctgaacct	atcgaggcca	acctctgtg	gcggcaggag	8520
	atgggcggga	acatcacccg	cgtggagtc	gaaaataagg	tagtaatttt	ggactctttc	8580
	gagccgctcc	aagcggagg	ggatgagagg	gaagtatccg	ttccggcgga	gatcctcgcg	8640
	aggctccagga	aattccctcg	agcgtgccc	atatgggcac	gcccggatta	caacctccca	8700
	ctgttagagt	cctggaagg	cccggaactac	gtccctccag	tggtacacgg	gtgtccattg	8760
20	ccgctgcca	aggccccctc	gataccacct	ccacggaggg	agaggacggg	tgctctgtca	8820
	gaatctaccg	tgtcttctgc	cttggcggag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tctgtcgccg	tgcacagcgg	cacggcaacg	gctctctctg	accagccctc	cgacgcagcg	8940
	gacgcgggat	ccgacgttga	ctcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggctc	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgtctgt	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgctggg	9120
25	gaaaccaagc	tgcccgctca	tgcaactgag	aactctttgc	tccgtcacca	caacttgggt	9180
	tatgtctaca	catctcgcat	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
	caggctctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aactctctatc	cgtggaggaa	gcctgttaag	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaggagc	gtccggaaac	tatccagcaa	ggccgttaac	9420
	cacatccgtc	ccgtgtggaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccacc	9480
30	atcatggcaa	aaaatgaggt	tttctgctgc	caaccagaga	agggggcgcg	caagccagct	9540
	cgcccttatcg	tattcccgaa	tttgggggtt	cgtgtgtgct	agaaaatggc	cctttacgat	9600
	gtggtctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tccagttcct	ggtgaatgcc	tgaaaagcga	agaaatgccc	tatgggcttc	9720
	gcatacgaca	cccgctgttt	tgactcaacg	gtcactgaga	atgacatccg	tggttgaggag	9780
35	tcaatctacc	aattgttgta	cttggccccc	gaagccagac	aggccataag	gtcgctcaca	9840
	gagcggtctt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggtctatcg	9900
	cggtgcccgc	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttatttg	9960
	aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgtatgcgga	10020
	gacgaccttg	tccgttatctg	tgaaagcgog	gggacccaag	aggacgaggg	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtcgcgca	cgatgcactt	10200
	ggcaaaaggg	tgtactatct	caccogtgac	cccaccaccc	cccttgccgg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcatttcc	tggttaggca	acatcatcat	gtatgcgccc	10320
	accttggtgg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctaacc	agatcattca	acgactccat	ggccttagcg	catttttact	ccatagttag	10500
45	tctccaggtg	agatcaatag	gggtggctta	tgccctcagg	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggg	tactgtccca	gggggggagg	10620
	gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccgg	ctgcgtccca	gttggaattta	tccagctggg	tccgttctgg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcggtgc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
	ctactttctg	taggggtagg	catctatcta	ctcccaacc	gatgaacggg	gagctaaaca	10860
50	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
	tttttttttt	tttttttttc	tctttttttt	ttctcttttt	tttctttttt	tttctttttg	10980
	tggctccatg	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtc	gatactggcc	tctctgcaga	tcaagt			11076

<210> 8
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 8
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctcgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtgggtctg cggaaccggg gagtacaccg gaattgccag 180
 gacgacccggg tcctttcttg gatcaaccog ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300
 gtgcttgcca gtgccccggg aggtctcgtg gaccgtgcac catgagcacg aatcctaacc 360
 ctcaaaagaaa aaccaaaccgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cagcagcggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctccctgcga gaaagtatcc atcatggctg atgcaatgcg gcggctgcac acgcttgatc 780
 cggctacctg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggagccggg tcttgcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
 ccgaactggt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
 ttgctgaaga gcttggcggc gaattggcgt accgcttctc cgtgctttac ggtatcgccg 1140
 ctccccgatc gcagcgcac gccttctatc gccttcttga cgagtcttct tgagttaaa 1200
 cagaccacaa cggtttccct ctacgcggat caattccgcc cctctccctc cccccccct 1260
 aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
 tcaccatatt tgcgtctttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggggtct ttccctctc gccaaaggaa tgcaaggctc gttgaatgtc 1440
 gtgaagggaag cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgaggcgagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac atgcaaaggc ggcacaaccc cagtgcacg ttgtgagttg gatagttgtg 1620
 gaaagagtca atggcctctc ctcaacgcta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccattt gtatgggac tgatctgggg cctcgggtga catgctttac atgtgtttag 1740
 tcgagggttaa aaaacgtcta gggcccccga accacgggga cgtgttttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcacatca ctacgtcac agggccggac aggaaccagg tcgaggggga ggtccaagtg 1920
 gtctccaccg caacacaatc ttctctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
 tatctggtg ccggctcaaa gaccttgcc ggcccaagg gccaatcac ccaaatgtac 2040
 accaatgttg accaggacct cgtcggctgg caagcgcgcc cgggggcgcg ttccttgaca 2100
 ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160
 gtgcgcgggc ggggcgacag cagggggagc ctactctccc ccaggcccg tctctacttg 2220
 aagggtctct cgggcggtcc actgctctgc cctcggggc acgctgtggg catcttctcg 2280
 gctgccgtgt gcaccgagg ggttgcgaag gcggtggact ttgtaccctg cgagtctatg 2340
 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacgcccc actggtagcg gcaagagcac taagggtgcc 2460
 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgtccgt cgcgcgccac 2520
 ctagggttctg gggcgatat gtctaaaggc catggtatcg accctaaccat cagaatcggg 2580
 gtaaggacca tcaccacggg tgccccatc acgtactcca cctatggcaa gtttcttgcc 2640
 gacggtggtt gctctggggg cgctatgac atcataatat gtgatgagtg ccaactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760
 cgaactcgtg tgcctgcac cgctacgct cgggatcgg tcaccgtgcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctcggcg cgaagctgtc cggcctcgga ctcaatgtg tagcatatta ccggggcctt 3000
 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgtctaatg 3060
 acgggtctta ccggtgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtcgacttca gccgggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
 gtgactccag gagaacggcc ctcgggcatt ttcgattcct cggttctgtg cgagtgtcat 3300
 gacgcgggct gtgcttggtg cgagctcacg cccgcggaga cctcagttag gttgcgggct 3360
 tacctaataa caccagggtt gcccgctg caggaccatc tggagttctg ggagagcgtc 3420
 ttacaggcc tcaccacat agacgcccac ttcttgtccc agactaagca ggcaggagac 3480
 aacttccct acctggtagc ataccaggct acggtgtgcg ccagggtcga ggctccacct 3540

55

	ccatcgtggg	accaaagtgt	gaagtgtctc	atacggctaa	agcctacgct	gcacggggcca	3600
	acgccccctgc	tgtataggct	gggagccgtt	caaaacgagg	ttactaccac	acaccccata	3660
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tcgtcacgag	cacctgggtg	3720
5	ctggtaggcg	gagtcctagc	agctctggcc	gcgtattgcc	tgacaacagg	cagcgtggtc	3780
	atttgtgggca	ggatcatctt	gtccggaaag	ccggccatca	ttcccagacag	ggaagtcctt	3840
	taccgggagt	tcgatgagat	ggaagagtgc	gcctcacacc	tcctttacat	cgaacaggga	3900
	atgcagctcg	ccgaacaatt	caaacagaa	gcaatcgggt	tgctgcaaac	agccaccaag	3960
	caagcggagg	ctgctgtctc	cgtggtggaa	tccaagtggc	ggaccatcga	agccttctcg	4020
	gcgaagcata	tgtggaatct	catcagcggg	atacaatatt	tagcaggctt	gtccactctg	4080
	cctggcaacc	ccggatagc	atcactgatg	gcattcacag	cctctatcac	cagcccgtctc	4140
10	accacccaac	ataccctcct	gtttaacatc	ctggggggat	gggtggccgc	ccaacttgct	4200
	cctccagcg	ctgcttctgc	tttcgtaggc	gccggcatcg	ctggagcggc	tgttggcagc	4260
	atagggccttg	ggaaggtgct	tgtggatatt	ttggcagggt	atggagcagg	ggtggcaggc	4320
	gcgctcgttg	cctttaaggt	catgagcggc	gagatgccct	ccaccgagga	cctggttaac	4380
	ctaactccctg	ctatectctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctgctgtcggc	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
15	ttcgtctcgc	ggggtaaaca	cgtctccccc	acgcactatg	tgccctgagag	cgaagcgtga	4560
	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgctgaa	gaggcttcac	4620
	cagtggatca	acgaggactg	ctccacgcca	tgctccggct	cgtggctaag	agatgtttgg	4680
	gattggatat	gcacgggtgtt	gactgatttc	aagacctggc	tccagtccaa	gctcctgccg	4740
	cgtattgcgg	gagtcctcct	cttctcatgt	caacgtgggt	acaagggagt	ctggcggggc	4800
	gacggcatca	tgcaaacac	ctgcccattg	ggggcacaga	tcaccggaca	tgtgaaaaac	4860
	ggttccatga	ggatcgtggg	gcctaggacc	tgtagtaaca	cgtggcatgg	aacattcccc	4920
20	attaacgcgt	acaccacggg	ccctctgcag	ccctcccccg	cgccaaatta	ttctaggggc	4980
	ctgtggcggg	tggtctgtga	ggagtacgtg	gaggttacgc	gggtggggga	tttccactac	5040
	gtgacgggca	tgaccactga	cgacgtaaag	tgcccggtgc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tggtgggggt	gcggttgac	aggtacgctc	cagcgtgcaa	acccctccta	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatcaatacc	tggttgggtc	acagctcccc	5220
	tgcgagcccg	aaccggatgt	agcagtgtct	acttccatgc	tcacogacce	ctccacattc	5280
25	acggcggaga	cggctaagcg	taggtcggcc	aggggatctc	ctccccctt	ggccagctca	5340
	tcagctagcc	agctgtctgc	gccttccctg	aaggcaacat	gcactaccgg	tcagtactcc	5400
	ccgacgcctg	acctcatcga	ggccaacctc	ctgtggcggc	aggagatggg	cgggaacatc	5460
	accgcgctgg	agtcagaaaa	taaggtagta	atthtggact	ctttcgagcc	gctccaagcg	5520
	gaggaggatg	agagggaagt	atccgttccg	gcggagatcc	tgccggaggtc	caggaaattc	5580
	cctcgagcga	tgcccatatg	ggcacgcgcg	gattacaacc	ctccactgtt	agagtcctgg	5640
	aaggaccggg	actacgtccc	tcagtggtga	cacgggtgtc	cattgccggc	tgccaaaggc	5700
30	cctccgatac	caccttcacg	gaggaagagg	acggttgtcc	tgtcagaate	taccgtgtct	5760
	tctgccttgg	cggagctcgc	cacagagacc	ttcggcagct	ccgaatcgtc	ggccgtcgac	5820
	agcggcacgg	caacggcctc	tcctgaccag	ccctccgacg	acggcgacgc	gggatccgac	5880
	gttgagtctg	actcctccat	gccccccctt	gagggggagc	cgggggatcc	cgatctcagc	5940
	gacgggtctt	ggtctaccgt	aagcgaggag	gctagttagg	acgtcgtctg	ctgctcgatg	6000
	tcctacacat	ggacaggcgc	cctgatcagc	ccatgcgctg	cggaggaaac	caagctgcc	6060
35	atcaatgcac	tgagcaacte	tttgctccgt	caccacaact	tggtctatgc	tacaacatct	6120
	cgcagcgcaa	gcctgcggca	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
	cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaaactt	6240
	ctatccgtgg	aggaagcctg	taagctgacg	ccccacatt	cggccagatc	taaatttggc	6300
	tatggggcaa	aggacgtccg	gaacctatcc	agcaaggccg	ttaaaccacat	ccgctccgtg	6360
	tggaaaggact	tgctggaaga	cactgagaca	ccaattgaca	ccaccatcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaaac	agagaagggg	ggccgcaagc	cagctcgcct	tatcgtattc	6480
40	ccagatttgg	gggttcgtgt	gtgcgagaaa	atggcccttt	acgatgtggt	ctccaccctc	6540
	cctcaggccg	tgatgggctc	ttcatacgga	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctggtga	atgcctggaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacacccgc	6660
	tggtttgact	caacgggtcac	tgagaatgac	atccgtgttg	aggagtcaat	ctaccaatgt	6720
	tgtgacttgg	cccccgaaagc	cagacaggcc	ataaggtcgc	tcacagagcg	gctttacatc	6780
	ggggggcccc	tgactaatc	taaagggcag	aactgcggct	atcgcgggtg	ccgcgcgagc	6840
45	ggtgtactga	cgaccagctg	cggttaatacc	ctcacatgtt	acttgaaggc	cgtgcggccc	6900
	tgctcgagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgtcgtt	6960
	atctgtgaaa	gcgcggggac	ccaagaggac	gaggcgagcc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	cctgggggac	ccgccccaaac	cagaatacga	cttggagttg	7080
	ataacatcat	gctcctccaa	tgtgtcagtc	gcgcacgatg	catctggcaa	aagggtgtac	7140
	tatctcacc	gtgacccac	cacccccctt	gcgcgggctg	cgtgggagac	agctagacac	7200
	actccagtea	attcctggct	aggcaacatc	atcatgtatg	cgccacctt	gtgggcaagg	7260
50	atgactctga	tgactcattt	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacct	acctcagatc	7380
	attcaacgac	tccatggcct	tagcgcattt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgcct	caggaaactt	ggggtaaccg	ccttgcgagt	ctggagacat	7500

EP 1 043 399 A2

	cgggccagaa	gtgtccgcgc	taggctactg	tcccaggggg	ggagggctgc	cacttgtggc	7560
	aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620
	tcccagttgg	atztatccag	ctggttcgtt	gctgggttaca	gcgggggaga	catatatcac	7680
5	agcctgtctc	gtgcccgacc	ccgctgggtc	atgtgggtgcc	tactcctact	ttctgtaggg	7740
	gtaggcacat	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
	ccatcctggt	tttttccctt	tttttttttc	tttttttttt	tttttttttt	tttttttttt	7860
	ttttctcctt	tttttttcc	ctttttttcc	ttttctttcc	tttgggtggc	ccatcttagc	7920
	cctagtcacg	gctagctgtg	aaaggcccg	gagccgcttg	actgcagaga	gtgctgatac	7980
10	tggcctctct	gcagatcaag	t				8001

15

20

25

30

35

40

45

50

55

<210> 9
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

10

15

20

25

30

35

40

45

50

55

```

<400> 9
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctgtgcag cctccaggac 120
ccccctccc gggagagcca tagtggctcg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtagt gcctgatagg 300
gtccttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcagc aatcctaaac 360
ctcaaagaaa aaccaaaagt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaaggagact gctgctattg ggcaagtgc cggggcagga tctcctgtca tctcaccttg 720
ctcctgcoga gaaagtatcc atcatggctg atgcaatgcg gcggtgcac acgcttgatc 780
cggttaactg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcacagggg ctgcgcgacg 900
ccgaactggt cgccaggctc aaggcgcgca tgccgcagcg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcgcc gaatgggtcg accgcttctc cgtgtcttac ggtatcgccg 1140
ctcccatttc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200
cagaccacaa cgttttccct ctagcgggat caattccgcc cctctccctc cccccccctc 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tcaccatata tgcctgtctt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctagggtgtc ttccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaag cagtctctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaggc ggcaacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaagagtgca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcacagaag 1680
gtacccattt gtatgggata tgatctgggg cctcgggtga catgtcttac atgtgtttag 1740
tcgagggtta aaaacgtcta gggcccccga accacgggga cgtggtttcc ctttgaaaaa 1800
cagcataata ccatgggcac gaatcctaaa cctcaaaaga aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcgggtg agttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc ctagcggtcg 1980
caacctcgtg gaagcgacac acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccc ggtacccctg gccctctat ggcaatgagg gcttgggggt ggacggatgg 2100
ctcctgtcac cccgtggctc tcggcctagt tggggcccca cgaccccccg gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccctc acgtgcggct tcgcccgatc catggggtag 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccttggcgca tggcgtccgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaaatctg ccggttgctc ctttctatc 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtacctagt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccgggtgc gtgccctgcg ttcgggagaa caactcctcc 2520
cgctgctggg tagcgtcac tcccacgctc gcggccagga acgctagcgt ccccaactacg 2580
acgatacgac gccatgtcga tttgctcgtt gggcgggctg ctctctgctc cgtatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgcccagc tgttccactt ctcgctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgcgtgggat atggtggcgg gggccattg gggagtccta 2880
gcgggccttg cctactatc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttccg gcgttgacgg gggaaacctat gtgacagggg ggaagatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcacccggg tcatccaga aaatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaacctggg 3120
ttccttgcgt cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
gccagctgca gcccacatga cgcgttcgct cagggggtgg ggcccatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtgcggtatc 3300
gtaccccgcg cgcaggtgtg tggccagatg tactgcttca ccccaagccc tgtcgtgggtg 3360
gggacgacgg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaaacgcgc gccgcgcaa ggcactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaaagac gtgcgggggc ccccgtgta acatcggggg gatcggcaat 3540

```

EP 1 043 399 A2

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaaccltga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
 tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actaccata caggcttttg 3660
 cactacccct gcaactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgttgtta cctggaggac 3780
 agggacagat cagagcttag cccgctgtcg ctgtctacaa cggagtggca ggtattgccc 3840
 tgttccctca ccaccctacc ggcctctgtcc actggtttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
 gagtatgtcc tgttgcctct ccttcttctg gcggacgcgc gcgtctgtgc ctgcttgtgg 4020
 atgatgtgc tgatagctca agctgagccc gccctagaga acctggtggg cctcaacgcg 4080
 gcatccgtgg ccggggcgca tggcattctc tccttctctg tgttcttctg tgctgcctgg 4140
 tacatcaagg gcaggctggt cccctggggcg gcatatgccc totacggcgt atggccgcta 4200
 ctctgctcc tcctggcggt accaccacga gcatacgcca tggaccggga gatggcagca 4260
 tctgtcggag gcgcggtttt cgtaggctcg atactcttga ccttgtcacc gcactataag 4320
 ctgttccctc ctaggctcat atggtggtta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggateccccc cctcaacgtt cggggggggc gcgatgcctg catcctctc 4440
 acgtgcgcga tccacccaga gctaatcttt accatcacca aaatcttgct cgcatactc 4500
 ggtccactca tgggtgtcca ggctggtata accaaagtgc cgtacttctg gcgcgcacac 4560
 gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620
 gctctcatga agttggccgc actgacaggt acgtacgttt atgacctct caccocactg 4680
 cgggactggg cccacgcggg cctacgagac cttgcggtgg cagttgagcc cgtcgtcttc 4740
 tctgatattg agaccaaggt tataccctgg ggggcagaca ccgcggcgtg tggggacatc 4800
 atcttggggc tgcccgctc gcgccgcagg gggagggaga tacatctggg accggcagac 4860
 agccttgaag ggcagggtg gcgactcctc gcgcctatta cggcctactc ccaacagacg 4920
 cgaggccctac ttggctgcat catcactagc ctacacaggc gggacaggaa ccaggctcag 4980
 ggggagggtc aagtggctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
 gtgtgtttga ctgtctatca tgggtgcggc tcaaaagacc ttgccggccc aaagggccca 5100
 atcacccaaa tgtacaccaaa tgtggacca gacctcgtcg gctggcaagc gccccccggg 5160
 gcgcgttcc tgcacacatg cacctgcggc agctcggacc tttacttggt cacgaggcat 5220
 gccgatgtca ttccggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5280
 cccgtctcct acttgaaggg ctcttcgggc ggtccactgc tctgccccctc ggggcacgct 5340
 gtgggcatct ttccggctgc cgtgtgcacc cgagggggtg cgaaggcggt ggactttgta 5400
 cccgtcgaat ctatggaac cactatgcgg tccccggtct tcacggacaa ctcgtccct 5460
 ccggccgtac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520
 agcactaagg tgccggctgc gtatgcagcc caagggtata aggtgcttgt cctgaaccgg 5580
 tccgtgcgg ccaccctagg ttccggggcg tatatgtcta aggcacatgg tatcgacct 5640
 aacatcagaa tcggggttag gaccatcacc acgggtgccc ccatacagta ctccacctat 5700
 ggcaagtttc ttgccgacgg tgggtgtctc gggggcgcc atgacatcat aatatgtgat 5760
 gagtgcact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
 gagacggctg gacgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggctacc 5880
 gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
 ggcaaaagcca tcccatcga gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
 aagaagaaat gtgatgagct cgcgcggaag ctgtccggcc tggactcaa tgctgtagca 6060
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgtcgtagca 6120
 acggacgctc taatgacggg cttaaccggt gacttcgact cagtgtatcga ctgcaatata 6180
 tgtgtcacc agacagtcca ctccagctg gacccgacct tcaccattga gacgacgacc 6240
 gtgccacaag acgcggtgtc acgctcgag cggcgaggca ggaactggtag gggcaggatg 6300
 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcattgtcga ttccctcggt 6360
 ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420
 gttaggttga gggcttacct aaacacacca ggggtgccc tctgccagga ccatctggag 6480
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtccagact 6540
 aagcaggcag gagacaact cccctacctg gtacataacc aggtacgggt gtgcgccagg 6600
 gctcaggctc cactccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660
 acgctgcacg ggccaacgcc cctgtgtat aggtggggag ccgttcaaaa cgaggttact 6720
 accacacacc ccataacca atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
 acgagcacct ggtgtgtgt aggcggagtc ctacgagctc tggccgcgta ttgctgaca 6840
 acaggcagcg tggctattgt gggcaggatc atcttgtccg gaaagccggc catcattccc 6900
 gacagggaag tcccttaccc ggagtctgat gagatggaag agtgccctc acacctccct 6960
 tacatcgaa agggaatgc gctcggcgaa caattcaaac agaaggcaat cgggttgctg 7020
 caaacgccca ccaagcaaag ggaggctgct gctcccggtg tggaaatcaa gtggcggaac 7080
 atcgaagcct tctgggcgaa gcatatgtgg aatttcatca cgggataca atatttagca 7140
 ggcctgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
 atcaccagcc cgtcaccac ccaacatacc ctccgttta acatcctggg gggatgggtg 7260
 gccgcccaac ttgctcctcc cagcgtgctc tctgctttcg taggcgcgg catcgtgga 7320
 gcggctgttg gcagcatagg ccttgggaag gtgcttgtgg atattttggc aggttatgga 7380
 gcagggttgg caggcgctc cgtggccttt aaggtcatga gcggcgagat gccctccacc 7440
 gaggacctgg ttaacctact ccctgctatc ctctccctcg gcgcctagt cgtcggggtc 7500

	gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggtatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggt	aaccacgtct	ccccacgca	ctatgtgcct	7620
	gagagcgacg	ctgcagcacg	tgctactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagaggc	ttcaccagtg	gatcaacgag	gaactgtcca	cgccatgctc	cggtcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggggc	acagatcacc	7920
	ggacatgtga	aaaacgggtc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaacat	tccccattaa	cgctacacc	acgggcccct	gcacgcccct	cccggcgcca	8040
10	aattattcta	ggcgctgtg	gcgggtggct	gctgaggagt	acgtggagggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	ggcatgacc	actgacgacg	taaagtggcc	gtgtcagggt	8160
	ccggcccccg	aattcttcac	agaagtggat	gggtgtcggt	tgacacaggta	cgctccagcg	8220
	tgcaaacccc	tctacgggga	ggaggtcaca	ttcctgggtcg	ggctcaatca	atacctgggt	8280
	gggtcacagc	tcccatgcga	gccgaaccg	gatgtagcag	tgctcacttc	catgctcacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctctctcc	8400
15	cccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actccccgga	cgctgacctc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atgggcggga	acatcacccg	cgtaggagtc	gaaaataagg	tagtaatttt	ggactctttc	8580
	gagcgcgtcc	aagcggagga	ggatgagagg	gaagtatccg	ttccggcgga	gatcctgcgg	8640
	aggctccagga	aattccctcg	agcgtatccc	atatgggcac	gcccggatta	caacctccca	8700
	ctgttagagt	cctggaagga	cccggaactc	gtccctccag	tggtacacgg	gtgtccattg	8760
20	ccgcctgcca	aggccctcc	gataccacct	tcacggagga	agaggacggg	tgctctgtca	8820
	gaatctaccg	tgtcttctgc	cttggcgag	ctcgccacag	agaccttcgg	cagctccgaa	8880
	tgcctggccg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcgggat	ccgacgttga	gtcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatccccgatc	tcagcgacgg	gtcttggctc	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctctgtct	cgatgtccta	cacatggaca	ggcgccctga	tcacgccatg	cgctgcggag	9120
25	gaaaccaagc	tgcccatcaa	tgactgagc	aactctttgc	tcctgtacca	caacttggtc	9180
	tatgtctaca	catctcgca	cgcaagcctg	cggcagaaga	agggtacact	tgacagactg	9240
	cagggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgttaagc	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaggac	gtccggaaac	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttctgt	gaagacactg	agacaccaat	tgacaccacc	9480
30	atcatggcaa	aaaatgagg	ttctgtcgct	caaccagaga	aggggggccc	caagccagct	9540
	cgcccttatcg	tattcccaga	tttgggggtt	cgtgtgtgcg	agaaaatggc	cccttacgat	9600
	gtggtctcca	ccctccctca	ggcgtgtatg	ggctcttcat	acggattcca	atactctctc	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tggaagcgga	agaaatgccc	tatgggcttc	9720
	gcatacgaca	cccctgtttt	tgactcaacg	gtcactgaga	atgacatccg	tggtgaggag	9780
	tcaatctacc	aattgttgta	cttggccccc	gaagccagac	aggccataag	gtcgtccaca	9840
35	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggtatctgc	9900
	cggtgcccgc	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
	aaggccgctg	cgccctgtcg	agctgcgaag	ctccaggact	gcacgatgct	cgatatcgga	10020
	gacgaccttg	tcgttatctg	tgaaagcgcg	gggacccaag	aggacgaggc	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgtctc	tccaatgtgt	cagtcgcgca	cgatgcatct	10200
	ggcaaaaagg	tgtactatct	caccctgtac	cccaccacc	cccttgccgg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggttaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccctct	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctaccct	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagtta	10500
45	tctccaggty	agatcaatag	ggtaggctca	tgccctcagga	aacttggggg	accgcccttg	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
	gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatccccg	ctgcgtccca	gttggtatga	tccagctggg	tcgttgctgg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
	ctactttctg	taggggtagg	catctatcta	ctcccaacc	gatgaacggg	gagctaaaca	10860
50	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
	tttttttttt	tttttttttc	tctttttttt	tctctttttt	tttctttttt	tttctttttg	10980
	tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtc	gatactggcc	tctctgcaga	tcaagt			11076

<210> 10
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 10
 gccagccccc gattgggggc gacactccac catagatcac tccccgtgta ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgctcgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtggctcg cggaaccggg gagtacaccg gaattgccag 180
 gacgaccggg tcctttcttg gatcaaccog ctcaatgcct ggagatttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
 gtgcttgcca gtccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaagaaa aaccaaacgt aacaccaacg ggccgcgcat gattgaacaa gatggattgc 420
 acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt 540
 ttgtcaaagc cagacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
 cgtggctggc cgtggcgggc gttcccttgc cagctgtgct cgacgttgtc actgaagcgg 660
 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720
 ctccctgccg gaaagtatcc atcatggctg atgcaatgcg gcggtgtcat acgcttgatc 780
 cggctacact ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
 tggagccgg tcttgcgat caggatgatc tggacgaaga gcatcagggg ctgcgcaccg 900
 ccgaactggt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
 actgtggccg gctgggtgtg gcggacggct atcaggacat agcgttggct acccgtgata 1080
 ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
 ctcccgatcc gcagcgatc gccttctatc gccttcttga cgagtcttcc tgagtttaaa 1200
 cagaccacaa cggtttccct ctagcgggat caattccgca cctctccctc cccccccct 1260
 aacgttactg gccgaagccg cttggaataa ggccgggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcgctcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
 acgagcattc ctagggtctc ttccctctc gccaaaggaa tgcgaaggtc gttgaatgtc 1440
 gtgaagggaag cagttctctc ggaagcttct tgaagacaaa caacgtctgt agcggacctt 1500
 tgaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg gcacaaacc cagtgccacg ttgtgagttg gatagtgtg 1620
 gaagagtgca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
 gtacccattt gtatgggacg tgatctgggg cctcgggtgca catgctttac atgtgtttag 1740
 tcgagggttaa aaaacgtcta gggcccccga accacgggga cgtggttttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcatcatca ctagcctcac agggccggac aggaaccagg tcgaggggga ggtccaagtg 1920
 gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
 ctacatgtgt cgggtcaaaa gaccttgcg ggcccaagg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggacct cgtcggctgg caagcgcgcc ccggggcgcg ttccttgaca 2100
 ccatgcacct gcggcagctc ggacctttac ttggtcagca ggcattgccg tgtcattccg 2160
 gtgcgcgggc ggggcgacag cagggggagc ctactctccc ccaggcccg tctcctactg 2220
 aagggctctt cgggcggtcc actgctctgc cctcggggc acgctgtggg catcttctcg 2280
 gctgccgtgt gcacccgagg ggttgcaag gcggtggact ttgtaccggt cgagtctatg 2340
 ggaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tggcccatct acacgcccct actggtagcg gcaagagcac taagggtgcc 2460
 gctgcgtatg caggccaagg gtataagtg cttgtcctga acccgtcgt cgccgccacc 2520
 ctagggttct gggcgatat gtctaaggca catggtatcg accctaaccat cagaatcgcg 2580
 gtaaggacca tcaccacggg tgcccccac acgtactcca cctatggcaa gttcttgcc 2640
 gacgggtggt gctctgggg cgctatgac atcataatat gtgatgagt ccaactcaact 2700
 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagac ggctggagcg 2760
 cgactcgtcg tgcctgcac cgctacgcct ccgggatcgg tcaccggtcc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactgga gaaatccct tttatggcaa agccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
 gagctcggcg cgaagctgtc cggcctcgga ctcaatgtc tagcatatta ccggggcctt 3000
 gatgtatccg tcataccac tagcggagac gtcattgtcg tagcaacgga cgtctaatg 3060
 acgggtctta ccggcgactt cgactcagtg atcgactgca atacatgtgt caccagaca 3120
 gtcgacttca gcctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtctat 3300
 gacgcgggct gtgcttggtc cgagctcacg cccgcggaga cctcagttag gttgcgggct 3360
 tacctaatac caccagggtt gcccgctctg caggaccatc tggagttctg ggagagcgct 3420
 tttacaggcc tcacccacat agacgcccat ttcttgtccc agactaaagca ggcagagac 3480
 aacttcccc acctggtagc ataccaggct acggtgtgcg ccagggtcga ggctccacct 3540

55

EP 1 043 399 A2

5 ccacgtgtgg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacggggcca 3600
 acgccccctgc tgtataggct gggagccgtt caaacagagg ttactaccac acaccccata 3660
 accaaatata tcatggcatg catgtcggtt gacctggagg tcgtcacgag cacctgggtg 3720
 ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
 attgtgggca ggtatcatct gtccggaaag ccggccatca ttcccgacag ggaagtcctt 3840
 taccgggagt tcgatgagat ggaagagtgc gctcacacc tcccttacct cgaacaggga 3900
 atgcagctcg ccgaacaatt caaacagaag gcaatcggtt tgctgcaaac agccaccaag 3960
 caagcggagg ctgctgctcc cgtggtggaa tccaagtggc ggaccatcga agccttcttg 4020
 gcgaagcata tgtggaattt catcagcggg atacaattt tagcaggctt gtccactctg 4080
 cctggcaacc ccgcatagc atcactgatg gcattcacag cctctatcac cagcccgctc 4140
 10 accaccaac ataccctcct gtttaacatc ctgggggatg gggtgccgc ccaacttgct 4200
 cctccacagg ctgcttctgc tttcgtaggg gccggcatcg ctggagcggc tgttggcagc 4260
 ataggccttg ggaaggtgct tgtggatatt ttggcagggt atggagcagg ggtggcaggc 4320
 gcgctcgtgg cctttaaggt catgagcggc gaggttgcct ccaccgagga cctggttaac 4380
 ctactccctg ctatcccttc ccttggcgcc ctagtctgct gggctcgttg cgacgcgata 4440
 ctgctgctgg acgtggggcc aggggagggg gctgtgcagt ggtgaaccg gctgatagcg 4500
 15 ttccgtctgc ggggtaacca cgtctccccc acgcactatg tgctgagag cgacgctgca 4560
 gcacgagtcga ctcatgctct ctctagtctt accatcactc agctgctgaa gaggcttcac 4620
 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgttttg 4680
 gattggatat gcacgggtgtt gactgatttc aagacctggc tccagtccaa gctcctgcgc 4740
 cgattgcccgg gactccctct ctctcatgt caacgtgggt acaaggaggt ctggcggggc 4800
 gacggcatca tgcaaacacc ctgcccattg ggagcacaga tcaccggaca tgtgaaaaac 4860
 ggttccatga ggatcgtggg gcctagacc tgtagtaaca cgtggcatgg aacattcccc 4920
 20 attaacgcgt acaccacggg cccctgcacg cctcccccgg cgccaaatta ttctaggcgc 4980
 ctgtggcggg taggtgctga ggagtaactg gaggttacgc ggggtggggg tttccactac 5040
 gtgacgggca tgaccactga caacgtaaag tgcccgtgct aggttccggc ccccgaaattc 5100
 ttacacagaag tggatggggt gcggttgcaac aggtacgctc cagcgtgcaa acccctccta 5160
 cgggaggagg tcacattcct ggtcgggctc aatcaatacc tgggttgggtc acagctccca 5220
 tgcgagcctg aaccggatgt agcagtgcct acttccatgc tcaccgacc ctccacattc 5280
 25 acggcgggaga cggctaagcg taggctggcc aggggatctc cccccccctt ggccagctca 5340
 tcagctagcc agctgtctgc gccctccttg aaggcaacat gcactaccg tcatgactcc 5400
 ccgagcgtc acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc 5460
 acccgctgtg agtcagaaaa taaggtagta attttgact ctttcgagcc gctccaagcg 5520
 gaggaggatg agaggggagt atccgttccg gcggagatcc tgccggaggtc caggaaattc 5580
 cctcgagcga tgcccatatg ggcacgccc gattacaacc ctccactggt agagtccctg 5640
 aagaccggc actacgtccc tccagtgtga cacgggtgct cattgcccgc tgccaaaggc 5700
 30 cctccgatac cacccttcacg gaggaagagg acggttgctc tgtcagaatc taccgtgtct 5760
 tctgccttgg cggagctcgc cacagagacc ttccggcagc ccgaatcgtc ggccgtcgac 5820
 agcggcacgg caacggcctc tcccgaccag cctccgacgc acggcgacgc gggatccgac 5880
 gttgagtcgt actcctccat gccccccctt gagggggagc cgggggatcc cgatctcagc 5940
 gacgggtctt ggtctaccgt aagcgaggag gctagttagg acgtcgtctg ctgctcgatg 6000
 tctcatacat ggacaggcgc cctgatcacg ccatgcgctg cggaggaaac caagctgcc 6060
 35 atcaatgcac tgagcaactc tttgtccgt caccacaact tgggtctatg tacaacatct 6120
 cgcagcgcaa accctgcgca gaagaaggtc acccttgaca gactgcaggt cctggacgac 6180
 cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
 ctatccgtgg aggaagcctg taagctgacg cccccacatt cggccagatc taaatttggc 6300
 tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360
 tggaggact tgcgtggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
 40 gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgctt tatcgtattc 6480
 ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
 cctcaggccg tgatgggctc ttcatacgga ttccaatact ctccgtgaca ggggtcgag 6600
 ttcttggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacaccgc 6660
 tgttttgact caacgggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
 tgtgacttgg ccccggaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
 gggggccccc tgactaatte taaagggcag aactgcggct atcgccggtg ccgcccagc 6840
 45 ggtgtactga cgaccagctg cggttaatacc ctacatggtt acttgaaggc cgttcgggc 6900
 tgtcgagctg cgaagctcca ggactgcacg atgctcgat gcggagacga ccttgtcgtt 6960
 atctgtgaaa gcgcggggac ccaagaggac gaggcgagcc tacgggctt caccgaggct 7020
 atgactagat actctgcccc cctgggggac ccgcccaaac cagaatacga cttggagttg 7080
 ataacatcat gctcctccaa tgtgtcagtc gcgcacgatg catctggcaa aagggtgtac 7140
 tatctcacc gtgacccacc caccocctt gcgcgggctg cgtgggagac agctagacac 7200
 50 actccagtc aatcctggct aggcacacatc attctagctc agccacett gtgggcaagg 7260
 atgatcctga tgactcattt ctctccatc cttctagctc aggaacaaact tgaaaaagcc 7320
 ctgattgtgc agatctacgg ggccgtgtac tccattgagc cacttgacct acctcagatc 7380
 attcaacgac tccatggcct tagcgcattt tcactccata gttactctcc aggtgagatc 7440
 aatagggtgg ctctatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500

55

EP 1 043 399 A2

```

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
5 agcctgtctc gtgcccgacc ccgctgggtc atgtgggtgc tactcctact ttctgtaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt tttttccctt tttttttttt tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttcct ctttttttcc ttttctttcc tttgggtggc ccatcttagc 7920
cctagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
10 tggcctctct gcagatcaag t 8001

```

15

20

25

30

35

40

45

50

55

<210> 11
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

10

15

20

25

30

35

40

45

50

55

```

<400> 11
gccagccccc gattgggggc gacactccac catagatcac tccccgtgta ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggtctg cggaaccggg gactacaccg gaattgccag 180
gacgaccggg tcttttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaggccc ttgtggtact gcctgatagg 300
gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaaagaaa aaccaaacgt aacccaacg ggccgcgcat gattgaacaa gatggattgc 420
acgcagggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccctg 720
ctcctgcgca gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tctgtcgat caggatgatc tggacgaaga gcatcagggg ctgcgcgca 900
ccgaactgtt cgcagagctc aagcgcgcca tgcgcgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgcg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggcgg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg 1140
ctcccgatcc gcagcgcatc gccttctatc gccttcttga cgagtctctc tgagtttaaa 1200
cagaccacaa cgttttccct ctacggggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccgggtgt cgtttgtcta tatgttattt 1320
tccaccatat tgcgctcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggctc tccctctc gccaaggaa tgcaagggtc gttgaatgtc 1440
gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgaggcgagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaggc ggcaacaacc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tggccagaag 1680
gtacccattt gtatgggata tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgagggttaa aaaaagctta ggcgcccgga accacgggga cgtgggtttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggcagtcac gttcccgggc ggtggtcaga tcgtcgggtg agtttacctg 1920
ttgccgcgca gggggcccg gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaagcgacac acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccc ggtacctctg gccctctat ggcaatgagg gcttgggggt ggccaggtgg 2100
ctcctgtcac ccgctggctc tcggcctagt tggggcccca cgacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggctat cgataccctc acgtgcggct tcgccgatct catggggtag 2220
attccgctcg tcggcgcccc cctagggggc gctgccaggg ccttgccgca tggcgctcgg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgcct ctttctctac 2340
ttccttttgg ctttgcgtgc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtacctagt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
gcggacatga tcatgcatac ccccggtgct gtgccctgcg ttccggagaa caactcctcc 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt cccactacg 2580
acgatacgac gccatgtcga tttgctcggt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagatc tctgcggatc tgttttctc gtcgcccagc tgttcacett ctgcctcgc 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccaagt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgag 2820
ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggccattg gggagtccta 2880
gcgggccttg cctactatct catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctcttttcgg gcgttgacgg gggaaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcacccggg tcatccaga aaatccagct tgtaaacacc 3060
aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caaactggg 3120
ttccttgcgt cgtgttctta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180
gccagctgca gccccatcga cgcgttcgct caggggtggg ggcacatcac ttacaatgag 3240
tcacacagct cggaccagag gccttattgt tggcactacg caccocggcc gtcggtatc 3300
gtaccccgcg cgcaggtgtg tggtcagatg tactgcttca ccccaagccc tgcgtgggtg 3360
gggacgacgg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gcccgcgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt tcaccaagac gtgcgggggc ccccgctgta acatcggggg gatcgggcaat 3540

```

EP 1 043 399 A2

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaacettga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
 tgtgtgttcgg ggccttggtt gacacccaga tgcttggtcc actaccata caggcttttg 3660
 cactacccct gcactgtcaa ctttaccatc ttcaagggtta ggatgtacgt ggggggagtg 3720
 gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgttgtaa cctggaggac 3780
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtgga ggtattgcc 3840
 tgttctctca ccacccatcc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
 gtggacgtac aatacctgta cggtataggg tcggcgggtg tctcctttgc aatcaaatgg 3960
 gagtatgtcc tgttgctctt ccttctctg gcgacgcgc gcgtctgtgc ctgcttgtgg 4020
 atgatgctgc tgatagctca agctgaggcc gccctagaga acctggtggg cctcaacgcg 4080
 gcatccgtgg ccggggcgca tggattctc tcttctctg tgttctctg tgctgcctgg 4140
 tacatcaagg gcaggctggt cctggggcgc gcatatgcc tctacggcgt atggccgcta 4200
 ctctgctcc tgctggcgtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
 tcgtgcggag gcgcggtttt cgtaggtctg atactcttga ccttgtcacc gactataag 4320
 ctgttctccg ctaggctcat atggtgtgta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggaacccccc cctcaacgtt cggggggggc gcgatgccgt catcctctc 4440
 acgtgcgcga tccaccaga gctaattctt accatcacca aaatcttgct cgccatactc 4500
 ggtccactca tgggtctcca ggctggtata accaaagtgc cgtacttctg gcgcgcacac 4560
 gggctcatcc gtgcgatcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620
 gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccacactg 4680
 cgggactggg cccacgcggg cctacgagac cttgcgggtg cagttgagcc cgtcgtcttc 4740
 tctgatatgg agaccaaggt tatcactgg ggggcagaca ccgcgcgctg tggggacatc 4800
 atcttggggc tgcccgctcc cgcgcgcagg gggagggaga tacatctggg accggcagac 4860
 agccttgaag ggcaggggtg gcgactctc gcgcctatta cggcctactc ccaacagacg 4920
 cgaggccctac ttgctgcat catcactagc ctacacagcc gggacaggaa ccaggtcgag 4980
 ggggaggtcc aagtggctcc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
 gtgtgttgga ctgtctatca tggtgccggc tcaagacccc ttgccggccc aaagggccca 5100
 atcacccaaa tgtacacca tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
 gcgcgttctc tgacacatcg cactcgcgc agctcggacc tttacttggt cagcaggcat 5220
 gccgatgtca ttccgggtgc cggcggggc gacagcaggg ggagcctact cccccagg 5280
 cccgtctctc acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcacgct 5340
 gtgggcatct ttccgggtgc cgtgtgcacc cgagggttg cgaaggcggg ggactttgta 5400
 ccgctcgagt ctatgggaac cactatgcgg tcccgggtct tcacggacaa ctctctccct 5460
 ccggcgtagc cgcagacatt ccaggtggcc catctacacg cccctactgg tagcggcaag 5520
 agcactaagg tgccgggtgc gtatgcaggc caagggtata aggtgcttgt cctgaacccg 5580
 tccgtcgcgg ccaccctagg tttcggggcg tataatgtcta aggcacatgg tatcgacct 5640
 aacatcagaa tcggggttaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
 ggcaagtttc ttgcgcagcg tgggtgctct gggggcgccct atgacatcat aatatgtgat 5760
 gagtgcact caactgactc gaccactatc ctgggcatcg gcacagtcct ggaccaagcg 5820
 gagcggctg gagcgcgact cgtcgtgctc gccacgccta cgcctccggg atcgttcacc 5880
 gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
 ggcaagcca tcccatcga gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
 aagaagaaat gtgatgagct cgcgcggaag ctgtccggcc tcggactcaa tgtcttagca 6060
 tattaccggg gccttgatgt atccgtcata ccaactagcg gagacgtcat tgtcttagca 6120
 acggacgctc taatgacggg cttaaccggc gacttcgact cagtgtatcga ctgcaataca 6180
 tgtgtcacc agacagtcga cttcagctg gaccgcacct tcaccattga gacgacgacc 6240
 gtgccacaag acgcggtgtc acgctcgag cggcgaggca ggactggtag gggcaggatg 6300
 ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcattgtcga ttctcgtgt 6360
 ctgtgcgagt gctatgacgc gggctgtgct tggtagagc tcacgcccgc cgagacctca 6420
 gttaggttgc gggcttacct aaacacacca ggggtgccc tctgccagga ccatctggag 6480
 ttctgggaga gcgtctttac aggcctcacc cacatagacg cccattttct gtcccagact 6540
 aagcaggcag gagacaaatt cccctacctg gtacataacc aggctacggt gtgcgccagg 6600
 gctcctcatc cactccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660
 acgctgcacg ggcacaagcc cctgctgtat aggtgggag ccgttcaaaa cgaggttact 6720
 accacacacc ccataacca atacatcatg gcattgcatt cgggtgacct ggaggtcgtc 6780
 acgagcacct ggggtgctgt aggcggagtc ctacgagctc tggccgcgta ttgctgaca 6840
 acaggcagcg tggctattgt gggcaggatc atcttgctcg gaaagccggc catcattccc 6900
 gacagggaag tctttaccg ggagtctgat gagatggaag agtgcgcctc acacctccct 6960
 tacatcgaa agggaaatga gctcgcgaa caattcaaac agaaggcaat cgggttctg 7020
 caaacagcca ccaagcaagc ggaggctgct gctcccggtg tggaaatcaa gtggcgga 7080
 atcgaggcct tctgggcgaa gcatatgtgg aatttcatca cggggataca atatttagca 7140
 ggcttgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
 atcaccagcc cgtcaccac ccaacatacc ctctgttta acatcctggg gggatgggtg 7260
 gcgcgccaac ttgtctctcc cagcgtgctc tctgctttcg taggcgcgg catcgtgga 7320
 gcggtctgtg gcagcatagg ccttgggaag gtgcttgtg atattttggc aggttatgga 7380
 gcagggtgg caggcgcgtc cgtggccttt aaggtcatga cggcgagtt gccctccacc 7440
 gaggacctgg ttaacctact cctgctatc ctctccctg gcgcctagt cgtcggggtc 7500

	gtgtgcgcag	cgatactgoc	tcggcacctg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggg	aaccacgtct	ccccacgca	ctatgtgcct	7620
	gagagcgacg	ctgcagcacg	agtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagaggc	ttcacacgtg	gatcaacgag	gactgctcca	cggcatgctc	cggctcgtgg	7740
	ctaagagatg	tttgggattg	gatatgcaog	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggccta	ggaacctgtg	taacacgtgg	7980
	catggaacat	tccccattaa	cgcgtacacc	acggggccct	gcacgccctc	cccgccgcca	8040
10	aattattcta	gggcgctgtg	gocgggtagt	gctgaggagt	acgtggagggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcatgacc	actgacaacg	taaaagtccc	gtgtcagggt	8160
	cggcccccgc	aattcttcac	agaagtggat	gggtgocgt	tgacacaggta	cgtccagcgt	8220
	tgcaaaacccc	tcctacggga	ggaggtcaca	ttcctggctg	ggctcaatca	atacctgggt	8280
	gggtcacagc	tccccatgca	gcctgaaccg	gatgtagcag	tgctcacttc	catgetcacc	8340
	gaccctcccc	acattacggc	ggagacggct	aagcgtaggc	tggccagggg	atctcccccc	8400
15	cccttggcga	gctcatcagc	tagccagctt	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actccccgga	cgtgaccttc	atcgaggcca	acctcctgtg	gcggcaggag	8520
	atgggcggga	acatcacccg	cgtggagtca	gaaaataagg	tagtaattht	ggactctttc	8580
	gagccgctcc	aagcggaggga	ggatgagagg	ggagtatccg	ttccggcgga	gatcctgcgg	8640
	aggtccaggga	aattccctcg	agcgtatgcc	atatgggcac	gcccggatta	caacctccca	8700
	ctgttagagt	cctggaaggga	cccgacttac	gtccctccag	tggtagacgg	gtgtccattg	8760
20	ccgcctgcca	aggccctccc	gataccacct	tcacggagga	agaggacggg	tgctcctgtc	8820
	gaatctaccg	tgtcttctgc	cttggcgga	ctgcgccag	agacctcggg	cagctccgaa	8880
	tcgtcggcgg	tcgacagcgg	cacggcaacg	gcctctcctg	accagccctc	cgacgacggc	8940
	gacgcggggt	ccgacgttga	ctcgtactcc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggtct	accgtaagcg	aggaggctag	tgaggacgtc	9060
	gtctgtgtct	cgatgtccca	cacatggaca	ggcgccctga	tcacgccatg	cgtgcgggag	9120
25	gaaaccaagc	tgcccatcaa	tgactgagc	aactctttgc	tccgtcacca	caacttggtc	9180
	tatgctacaa	catctcgagc	cgcaaacctg	cggcagaaga	aggtcacctt	tgacagactg	9240
	cagggtcctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaaag	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaggac	gtccggaacc	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgcgt	gaagacactg	agacaccaat	tgacaccacc	9480
30	atcatggcaa	aaaatgaggt	tttctgcgtc	caaccagaga	aggggggccc	caagccagct	9540
	cgcccttatcg	tattcccaga	tttgggggtt	cgtgtgtgcg	agaaaatggc	cctttacgat	9600
	gtggtctcca	ccctccctca	ggccgtgatg	ggctcttcat	acggattcca	atactctcct	9660
	ggacagcggg	tcgagttcct	ggtgaatgcc	tgaaaagcga	agaaatgcc	tatgggcttc	9720
	gcatatgaca	cccgctgttt	tgactcaacg	gtcactgaga	atgacatccg	tgttgaggag	9780
	tcaatctacc	aatgttgtga	cttggccccc	gaagccagac	aggccataag	gtcgtccaca	9840
35	gagcgccttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cgggtgcgcg	cgagcgggtg	actgacgacc	agctgcggta	ataccctcac	atgttacttg	9960
	aaggccgctg	cggcctgtcg	ctccaggact	gcacgatgct	cgatgcgga	cgtatgcgga	10020
	gacgaccttg	tcgttatctg	tgaagcgcg	gggacccaag	aggacgaggg	gagcctacgg	10080
	gccttcacgg	aggctatgac	tagatactct	gccccccctg	gggacccgcc	caaaccagaa	10140
	tacgacttgg	agttgataac	atcatgctcc	tccaatgtgt	cagtgcgcga	cgatgcactc	10200
40	ggcaaaaggg	tgtactatct	caccctgtac	cccaccaccc	cccttgccgg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
	accttgtggg	caaggatgat	cctgatgact	catttcttct	ccatccttct	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctaccca	agatcattca	acgactccat	ggccttagcg	cattttcact	ccatagttac	10500
	tctccagggt	agatcaatag	ggtggcttca	tgcctcagga	aacttggggt	accgcccttg	10560
45	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctaggc	tactgtccca	gggggggagg	10620
	gctgccactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccg	ctcgcgtccc	gttggtttta	tcagctgggt	tcgttgctgg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttcatgtg	gtgcctactc	10800
	ctactttctg	taggggtagg	catctatcta	ctccccaacc	gatgaacggg	gagctaaaca	10860
	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
50	tttttttttt	tttttttttc	tccttttttt	ttcctctttt	tttccctttc	tttccctttg	10980
	tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtc	gatactggcc	tctctgcaga	tcaagt			11076

Patentansprüche

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten , dadurch gekennzeichnet,
5
daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.
- 10 2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,
daß die Hepatomazellen von einer handelsüblichen Hepatomazellinie abstammen .
3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,
15
daß die Hepatomazellen aus einer Hepatomaprimärzellkultur gewonnen sind.
4. HCV-RNA-Konstrukt, dadurch gekennzeichnet,
20
daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.
5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,
25
daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt.
6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,
30
daß die 3' NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT
TTAGCTTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTG
35
TTGCCCAGACTGGAGT

40

45

50

55

- 5 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GGAGC
- 10 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GCAGC
- 15 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT
- 20 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT
- 25 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT TTTTTTTTTT TTGAGAGAGA
GAGTCTCACT CTGTTGCCA GACTGGAGT
- 30 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGA CTGGAGT
- 35 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTTTAAT CTTTTTTTTT TTTTCCTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 40 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT
TTTTTAATC TTTTTTTTTT TTTCTTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

- 45 7. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.
- 50 8. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Neomycinphosphotransferasegen ist.
- 55 9. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet,
daß das selektierbare Markergen strangabwärts der 5' NTR in die HCV-RNA integriert ist.
10. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 9, dadurch gekennzeichnet,

daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.

- 5 11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet,
daß es ein integriertes Reportergen aufweist.
12. HCV-RNA-Konstrukt nach Anspruch 11, dadurch gekennzeichnet,
10 daß das Reportergen ein Gen aus der Gruppe der Luziferasogene, dem CAT-Gen (Chloramphenicol-Acetyl-
Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), der GFP-Gene (green-fluorescence-protein-Gene
) , dem GUS-Gen (Glukuronidasegen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen) ist.
13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 11, dadurch gekennzeichnet,
15 daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflusst.
14. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 13, dadurch gekennzeichnet,
20 daß das Resistenzgen derart in das offene Leseraster der HCV-RNA inkloniert ist, daß es erst nach einer pro-
teolytischen Prozessierung in eine aktive Form überführbar ist.
15. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 14, dadurch gekennzeichnet,
25 daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß
sie gemeinsam ein Fusionsprotein exprimieren.
16. Zellkultursystem nach einem der Ansprüche 1 bis 3 dadurch gekennzeichnet,
30 daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 15 ist.
17. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,
35 daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterle-
gungsnummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) hinterlegt sind.
18. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-
Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung und/oder Evaluierung und/oder Testung von The-
40 rapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.
19. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-
Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.
20. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung einer leberzellspezi-
45 fischen Genfahre für die Gentherapie .
21. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet,
50 daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzu-
schleusen, die zur Expression dieses Fremdgens geeignet ist.
22. Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts gemäß einem der
55 Ansprüche 4 bis 15, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz
aufweisen, dadurch gekennzeichnet,
daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial
ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 15 ist, auf/in dem dem Selektions-
gen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man

aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.

23. Verfahren nach Anspruch 22, dadurch gekennzeichnet,

5 daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 einschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

10

24. Verfahren zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV- Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, dadurch gekennzeichnet,

15 daß man mit einem Verfahren nach Anspruch 22 oder 23 eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt und isoliert,
daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt,
20 und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV- Vollängengenom oder ein HCV- Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

25. Zellkultur-adaptiertes HCV-RNA-Konstrukt mit hoher Replikationseffizienz, dadurch gekennzeichnet,

25

daß es durch Nukleotid- und/oder Aminosäure -Mutationen von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15 ableitbar ist und daß es mit einem Verfahren nach einem der Ansprüche 22 bis 24 erhältlich ist.

30 **26.** Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25, dadurch gekennzeichnet,

daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist.

35

27. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25 oder 26, dadurch gekennzeichnet,

daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.

40

28. Zellkultur-adaptierte Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, dadurch gekennzeichnet,

45 daß sie mit einem Verfahren erhältlich ist, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt nach Anspruch 24 durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 15, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

50

29. Hepatitis C Viruspartikel oder virus-ähnliche Partikel dadurch gekennzeichnet,

daß sie mit einem Verfahren nach einem der Ansprüche 22-24 erhältlich sind.

55 **30.** Zellen, infiziert mit Hepatitis C Viruspartikeln oder virus-ähnliche Partikeln gemäß Anspruch 29.

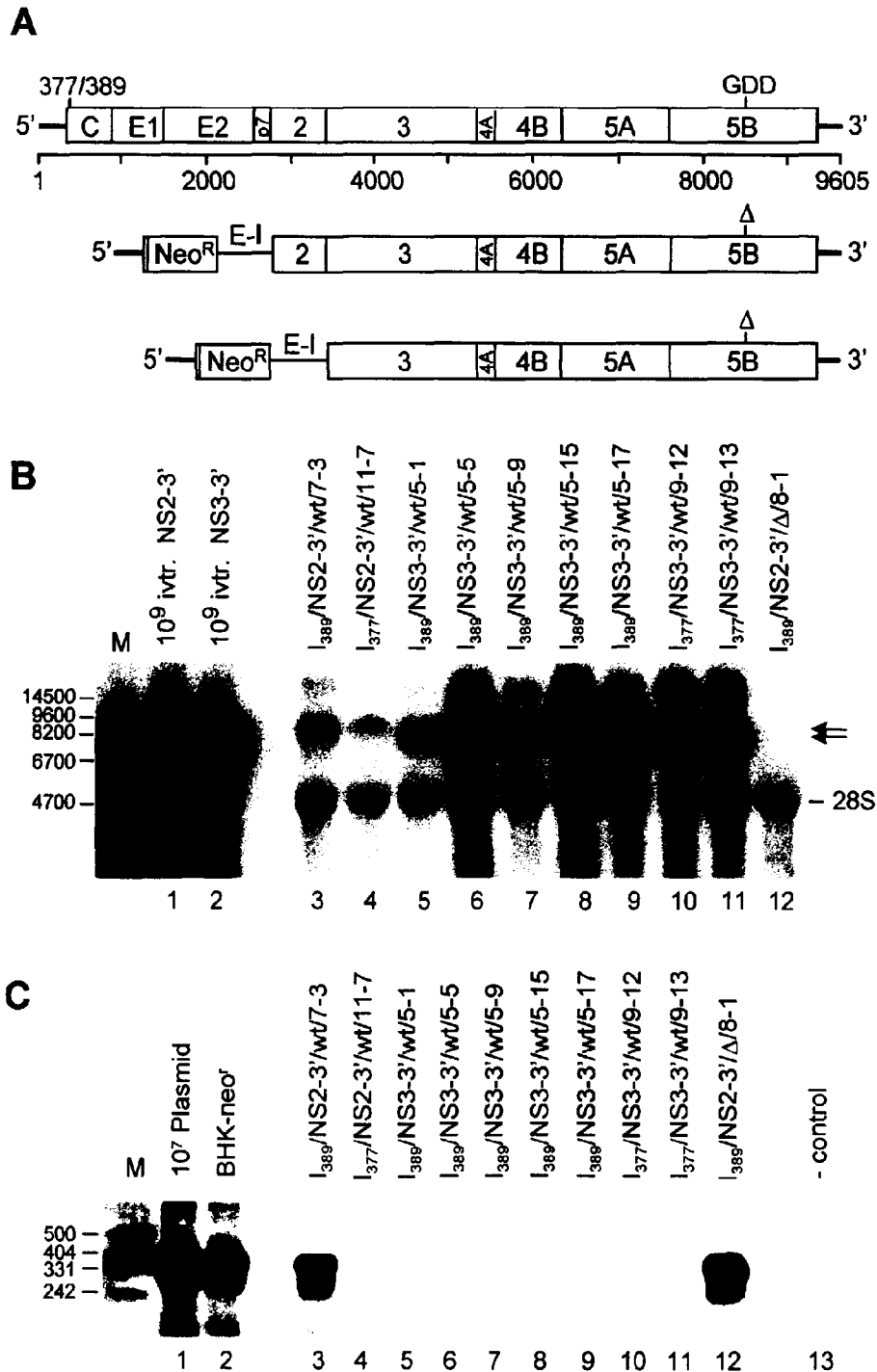


Fig. 1

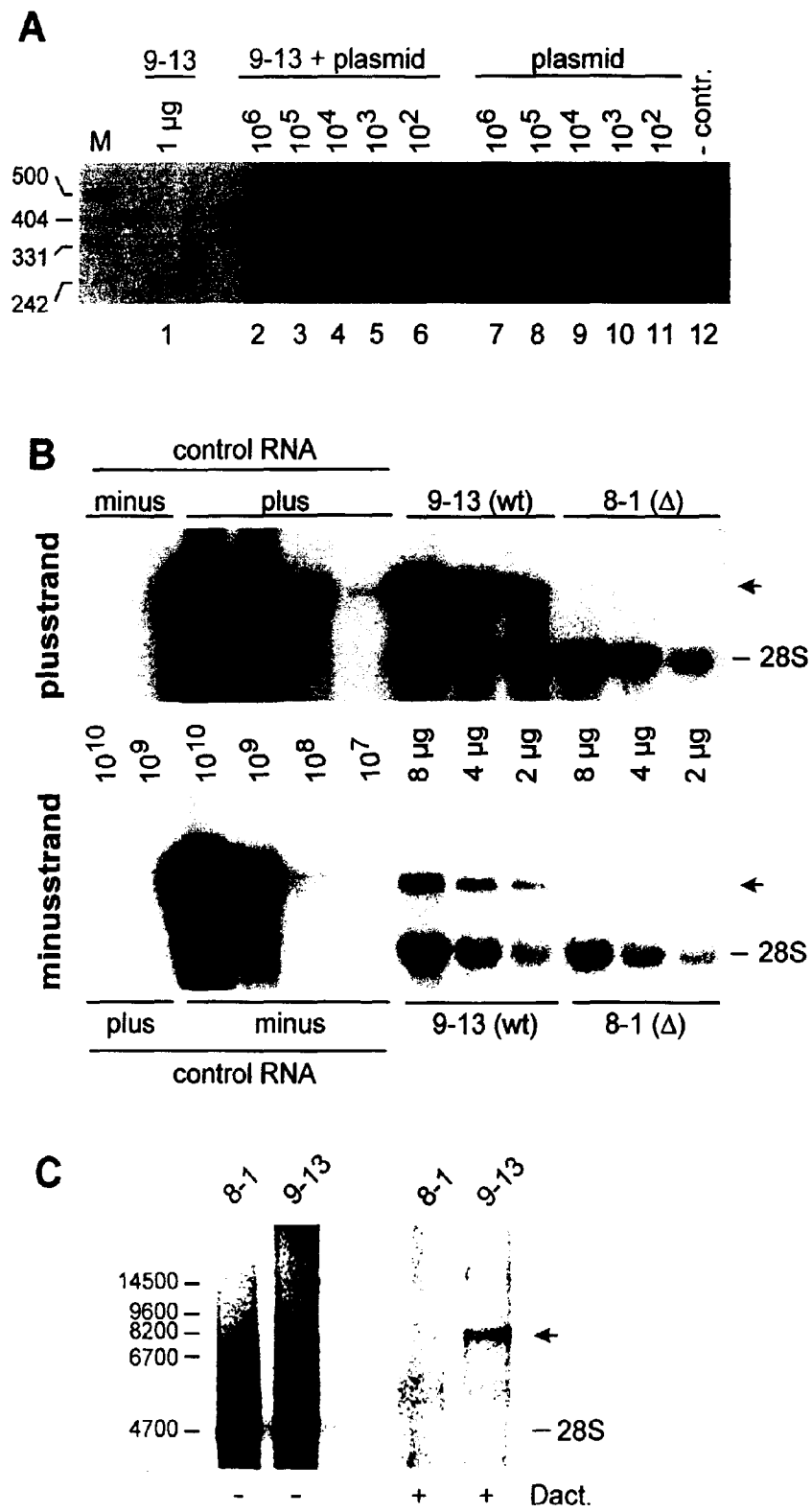


Fig. 2

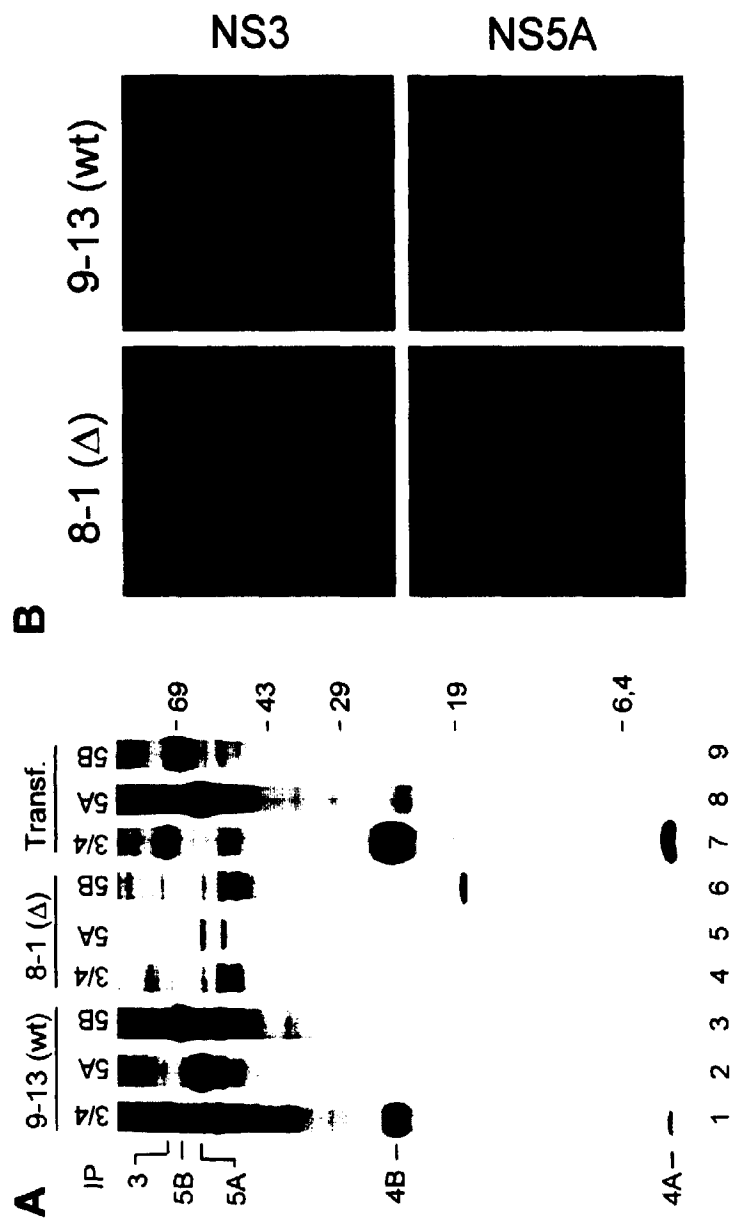


Fig. 3

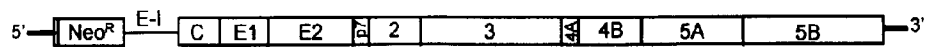


Fig. 4

A



B

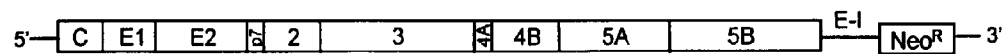
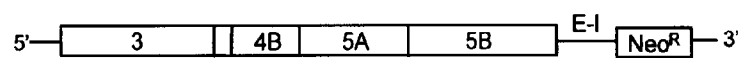


Fig. 5

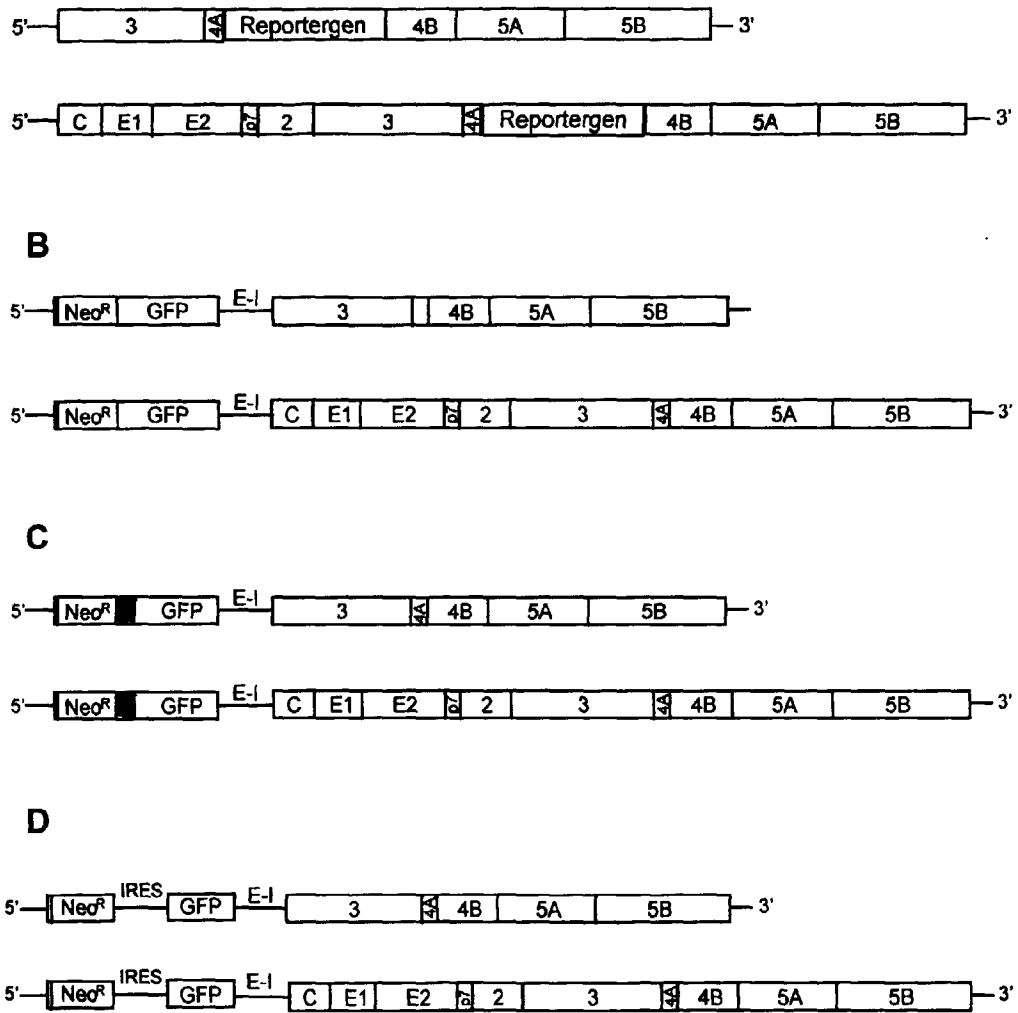


Fig. 6

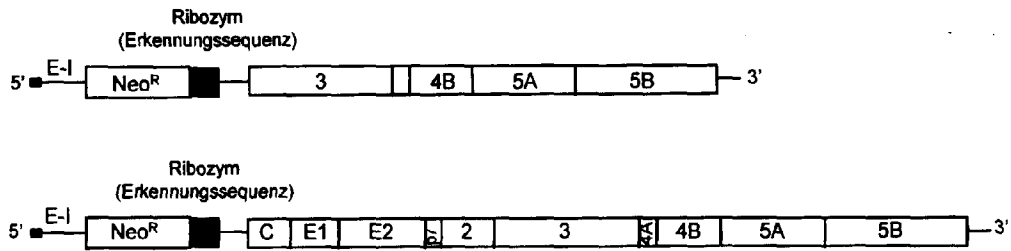


Fig. 7

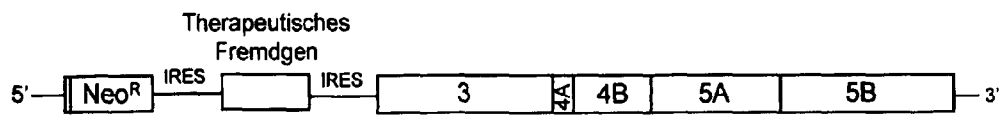


Fig. 8

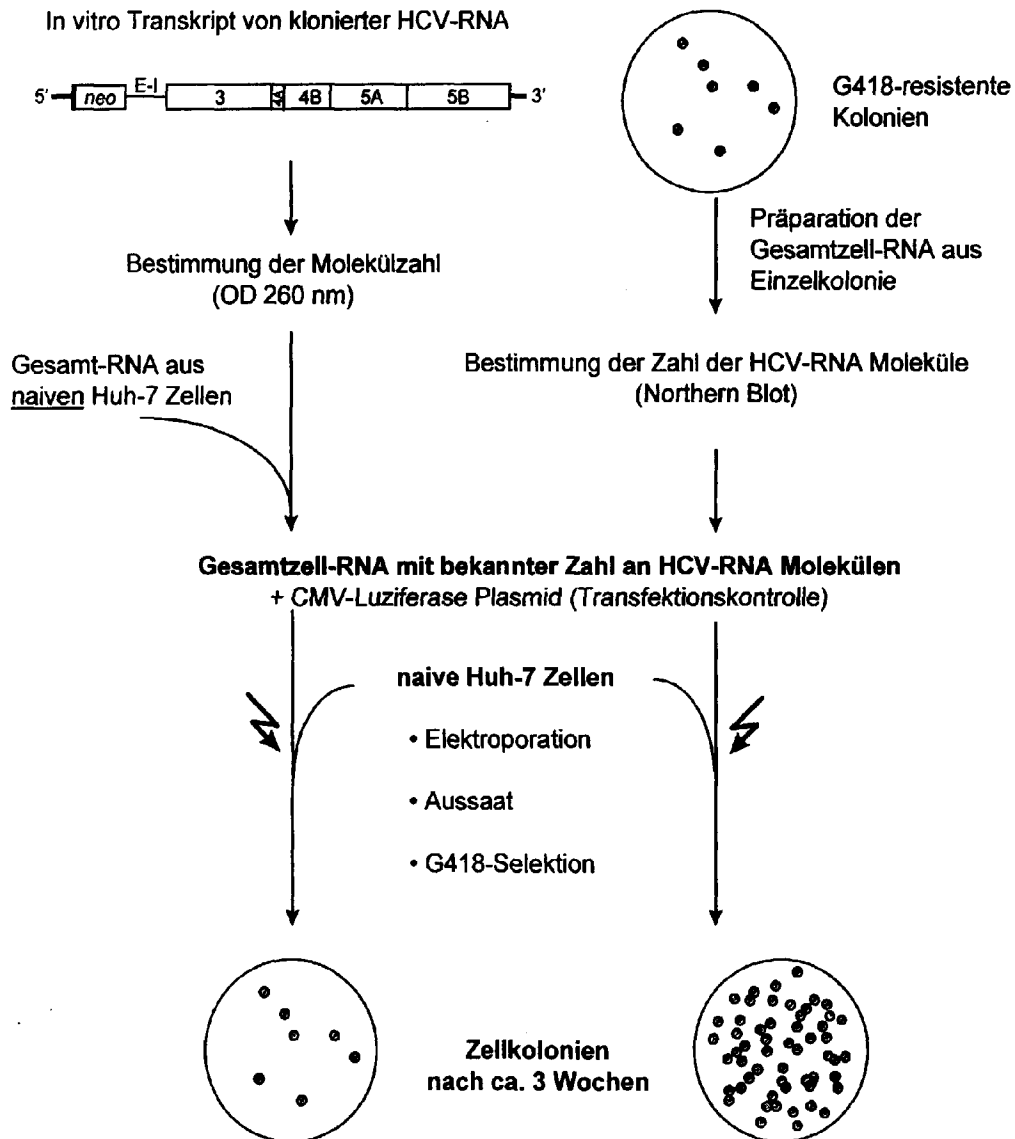


Fig. 9

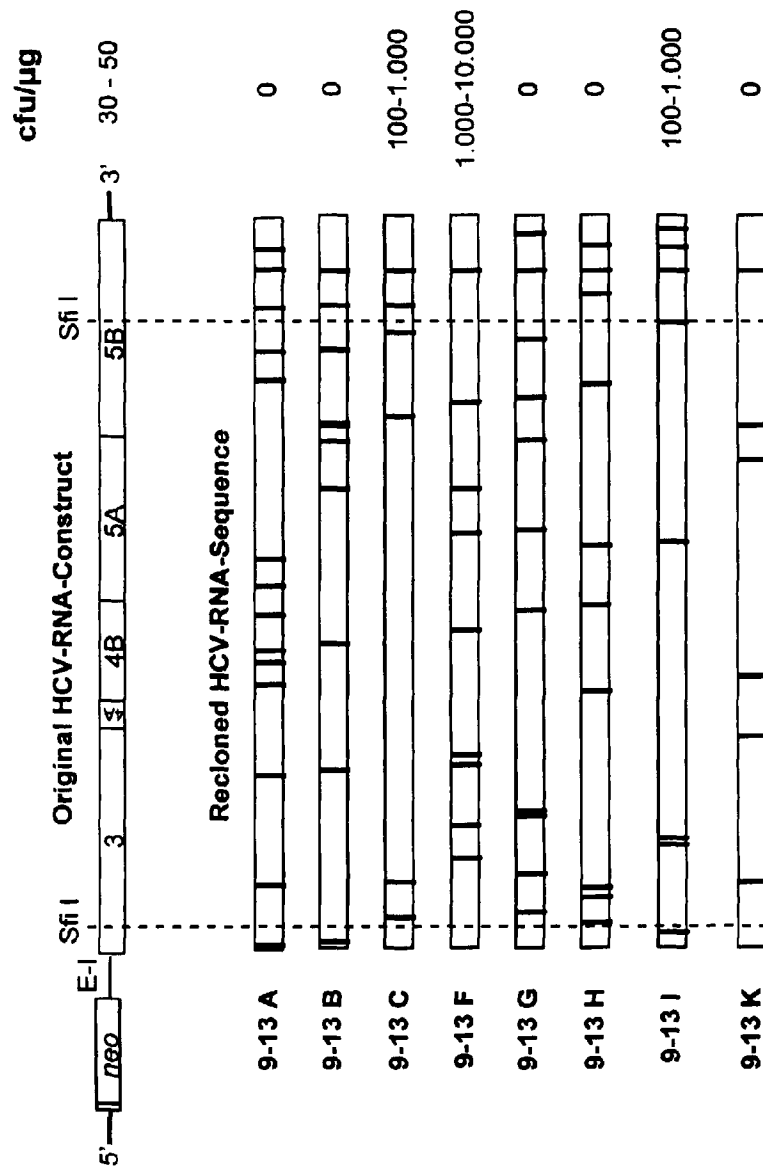


Fig. 10

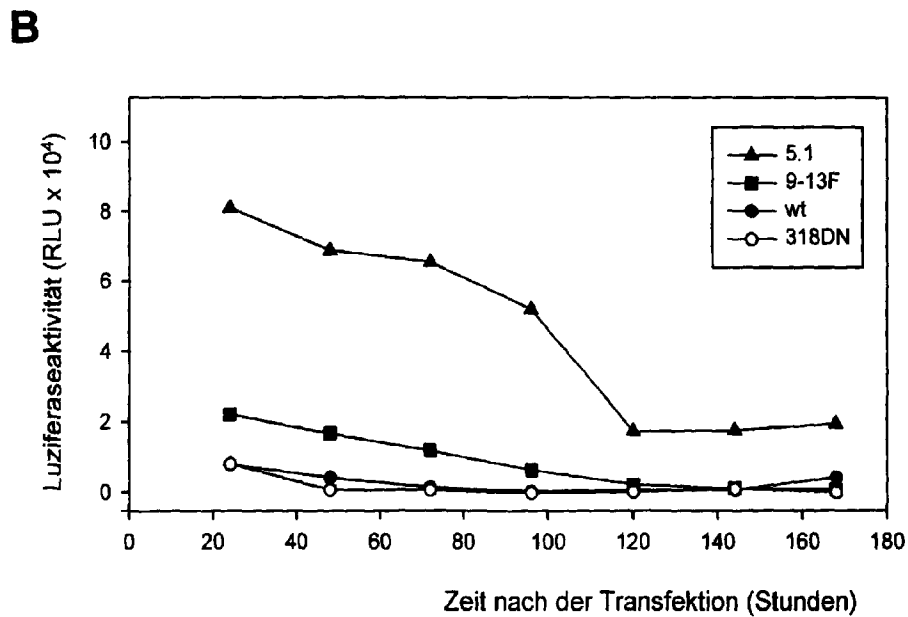
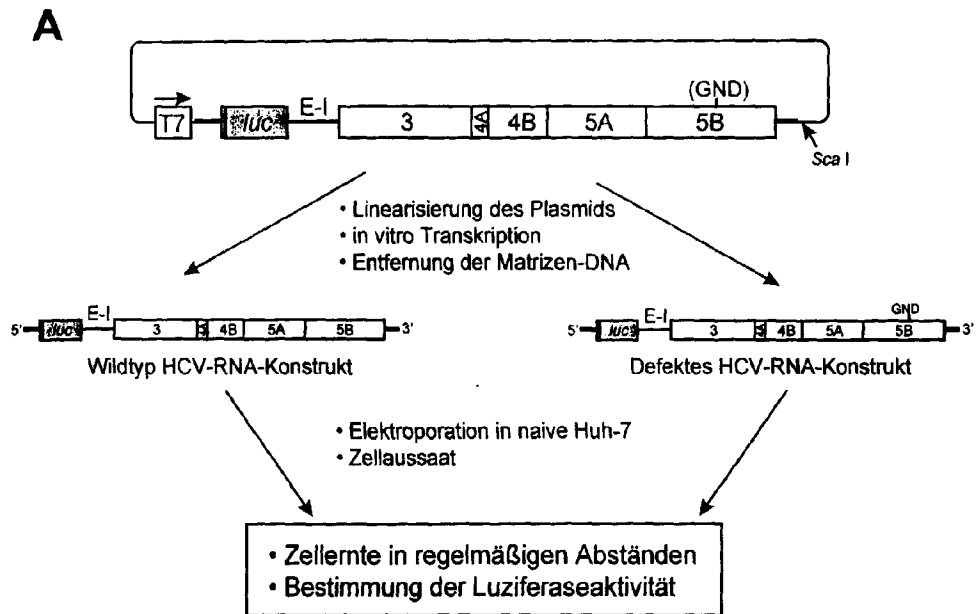


Fig. 11

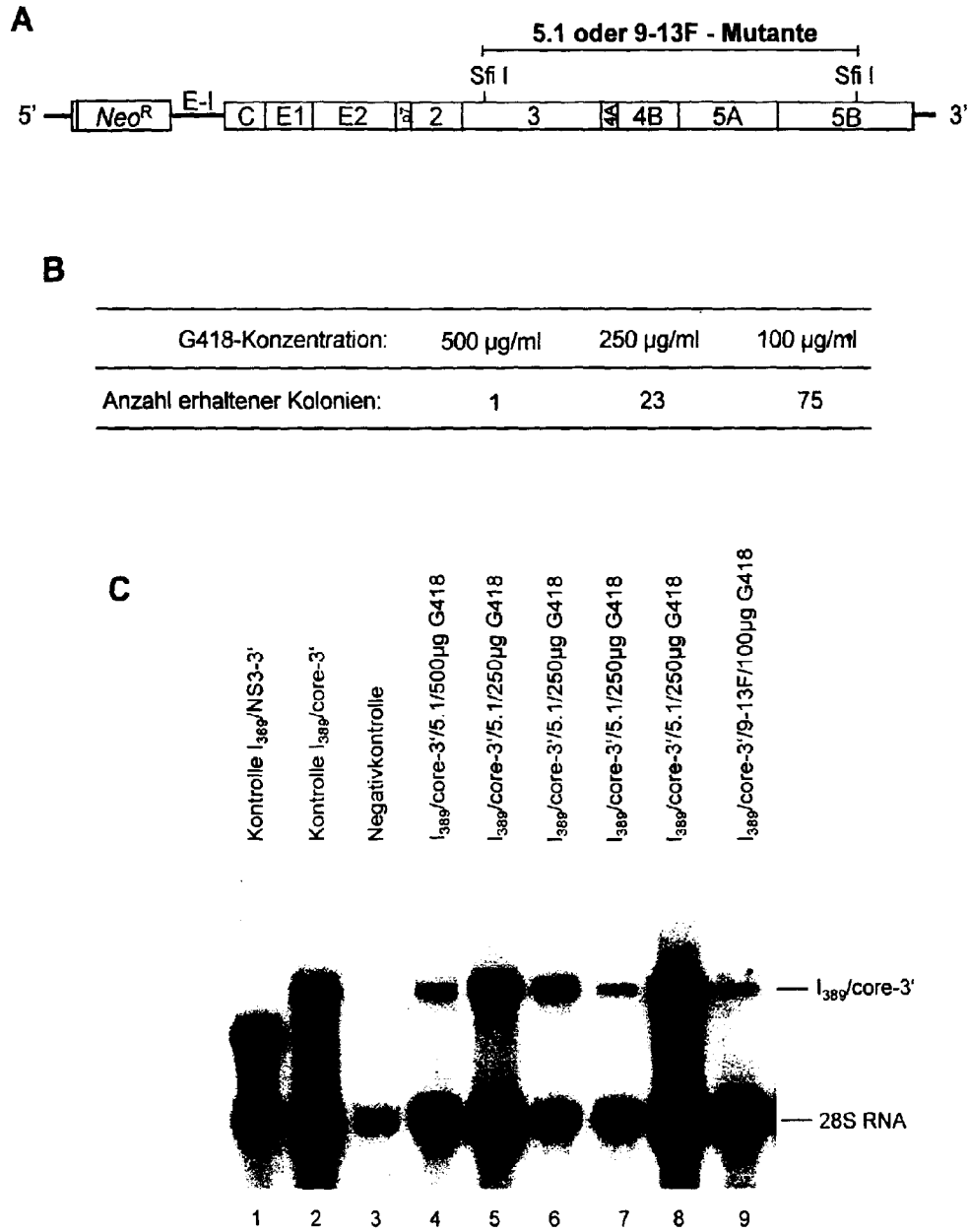


Fig. 12

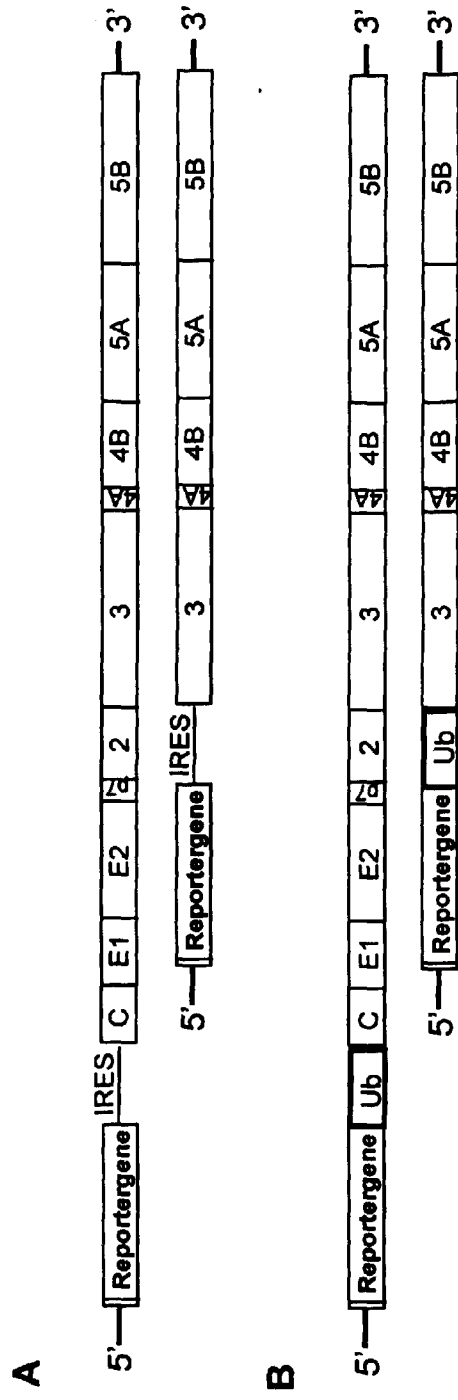


Fig. 13

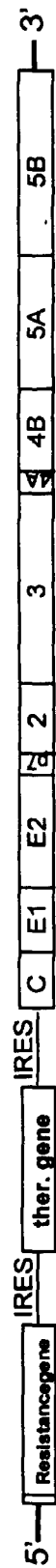


Fig. 14

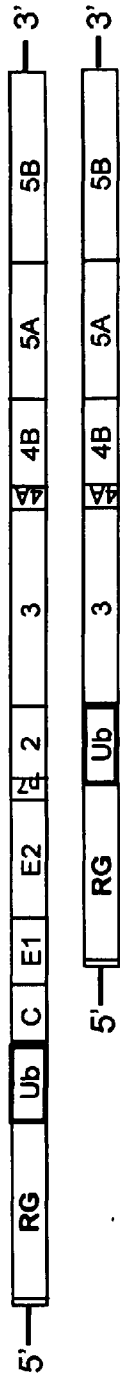


Fig. 15